2020

Essays on information revelation and political institutions

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

ESSAYS ON INFORMATION REVELATION AND POLITICAL INSTITUTIONS

by

MARIE THEA J. S. S. HOW CHOON

B.A., University of Cambridge, 2014

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2020

Approved by

First Reader

Dilip Mookherjee, PhD
Professor of Economics

Second Reader

Barton Lipman, PhD
Professor of Economics

Third Reader

Juan Ortner, PhD
Associate Professor of Economics

"Unless you are good at guessing, it is not much use being a detective."

Agatha Christie

Acknowledgments

First of all, I would like to thank Dilip Mookherjee for his patient and steady guidance during the past few years. Our conversations have been a lifeboat on a bottomless sea. I would also like to thank Bart Lipman and Juan Ortner for their thoughtful advice and impeccable teaching, without which much of my research would not have come to be.

I would like to thank my parents and my sisters for their truly unconditional support. It was a bittersweet fancy of mine, to follow an idea all the way to this far-flung edge of the world that is Boston. Thank you for accepting and nurturing my dreams. To Peter Tran, thank you for sitting with me through the storm.

Heartfelt thanks to Fatima Aqeel and Undral Byambadalai for years of friendship and goodness and believing in one another. I am grateful to all my friends, classmates and fellow BU students who have shared in this journey and brought me much joy.

I must thank Michael Manove, for teaching me about teaching; Larry Epstein, for teaching me to think; and Chiara Margaria, for being so incredibly cool. I am also grateful to many professors at BU and at Cambridge who have taught me a great deal about economics.

With regards to Chapters One and Two, I am indebted to participants of the BU Micro Theory Workshop and the International Conference on Game Theory at Stony Brook for their invaluable comments. I would also like to thank Stephen Morris, Marina Halac and Yingni Guo for insightful discussions.

With regards to Chapter Three, I am grateful to my co-authors, Jesse Bruhn and Anna Weber, for their dedication and enthusiasm. I am indebted to Raymond Fisman, Kevin Lang, Daniele Paserman and Lakshmi Iyer for their helpful comments and support, as well as the participants at the BU Empirical Micro Workshop, the BU Micro Theory Workshop and the Lecznar Summer Lectures for their excellent feedback. I would like to thank the Institute for Economic Development at Boston University for their generous financial support for this project, and to Hsueh-Ling Huynh and the Department of Economics for providing excellent research assistants. Many thanks to Lin Zhang, Wen Dao Xue, Yidi Wu, Rui Li, and Keshen Ling for their wonderful help.

All errors are my own.

ESSAYS ON INFORMATION REVELATION AND POLITICAL INSTITUTIONS

MARIE THEA J. S. S. HOW CHOON

Boston University, Graduate School of Arts and Sciences, 2020

Major Professor: Dilip Mookherjee, PhD
Professor of Economics

ABSTRACT

This dissertation examines how political institutions shape incentives for the transmission of policy-relevant private information. The first two chapters present theoretical frameworks to examine strategic information transmission to political candidates by a biased adviser. The third explores the causal impact of gender representation on municipal outcomes in the United States.

In Chapter One, I examine a model of "cheap talk" lobbying with no commitment. A biased adviser seeks to influence the policy outcome of a Downsian election by sending messages to the two candidates before they announce their policy platforms. I show that the adviser may credibly reveal some information on voter preferences, but only privately to one candidate. Political competition has a disciplining effect; the adviser prefers extreme policies, but instead recommends a pragmatic policy — one that is just close enough to voters' preference. In some situations, the presence of the biased adviser benefits the median voter.

The second chapter presents a model of informational lobbying with full commitment. The biased adviser strategically designs informative signals on voter preferences that will be observed by each of the two candidates. In contrast to the cheap talk context,

the optimal signal structure is shown to involve only public signals that are observed by both candidates. In particular, candidates receive precise information about how extreme voter preferences are, but not whether voters lean right or left. Consequently, both candidates choose the same biased policy, as a result of which the median voter is always worse off.

Chapter Three investigates the effect of gender representation on municipal outcomes in the United States between 2008 and 2016. Using novel data, the analysis exploits close elections between male and female candidates to measure the impact of an exogenous increase in the number of female council members. Consistent with the existing literature, we find evidence of decreased per capita expenditure, which, we argue, is not driven by revenue constraints but by increased disagreement or "gridlock" within the council. We also find no significant effect of gender representation on the composition of municipal spending or on other women's political aspirations.

Contents

List of Tables

List of Figures

List of Abbreviations

OLS	....	Ordinary Least Squares
RD	....	Regression Discontinuity

Chapter 1

Biased Campaign Advice: A Cheap Talk Model

1.1 Introduction

> "I immediately think of interest groups. That's how we gauge our public opinion.... I very rarely am clueless about where that constituency is because of the interest groups keeping me informed..."
>
> — Legislative staffer
>
> Source: (Herbst, 1998)

A substantial literature has been devoted to understanding political influence—how biased individuals or entities set about influencing policymakers' decisions. Such entities are sometimes referred to as special interests, and their activities as lobbying. Although models are wide-ranging, a large number have focused on strategic information transmission as a major form of lobbying.

Such models typically assume the special interest has superior knowledge of the optimal policy which the policymaker, as a conscientious social planner, should implement. This leads to a simple communication game where a biased sender (special interest) communicates with a receiver (policymaker) about an unknown state of the world (best policy), and the receiver takes an action that determines payoffs. However, this simplification is not without consequences. To the extent that policymakers are not social planners but elected politicians, and if some policy decisions must be announced before elections, the relevant information is not the best policy, but *voter preferences* on policy. Furthermore, the policy chosen by a candidate is not implemented unless they win. In that sense, policy decisions cannot be neatly disentangled from the election process.

In this chapter, I explicitly study how election incentives shape the role of special interests in providing — and manipulating — information on voter preferences.[1] The model is as follows. Two office-seeking candidates play a Downsian election game, where each simultaneously chooses a policy platform, which they commit to implementing if they win. Under majority voting, the policy platform preferred by the median voter must win. However, candidates are uncertain about voter preferences. A special interest (henceforth "sender") possesses this information and is able to send strategic messages to each candidate before they announce their policy platforms. The sender is extremely biased; she would like the implemented policy to be as high as possible. I show that this new model produces strikingly different results regarding the extent of information transmission and, importantly, the sender's ability to distort policy.

Suppose the sender cannot commit to a message strategy — that is cheap talk. In particular, she cannot promise not to lie, should the occasion arise. This creates a credibility problem: in the standard cheap talk model, extremely biased senders always lie, and therefore are never believed. In my model, however, political competition makes excessive lying disadvantageous. Suppose the sender only talks to one candidate. Then, if she falsely convinces that candidate to choose an unreasonably high policy, that policy will lose the election anyway, so it will not be implemented. Hence, the sender had best recommend a policy platform that is just appealing enough to the median voter. I show that an extremely biased sender is able to convey information, and even distort policy, in precisely this manner — giving private advice to one

[1] There is some evidence that politicians turn to interest groups for information about voter preferences ((Herbst, 1998)). Indeed, special interests tend to have the resources, incentive and comparative advantage in knowing how voters feel about particular issues.

candidate only. This prediction mirrors the kind of partnership that we observe in real life between politicians and their advisers. Under certain conditions, I show that even extremely biased advisers may, in fact, improve median voter welfare.

These welfare implications are of immediate and contemporary interest to observers of American politics. For instance, campaign advisers such as Stephen Bannon are widely thought to have been successful – at least to some extent – in promoting their personal policy agenda. The Heritage Foundation, a conservative think-tank, provides advice to a large number of politicians, with the explicit objective to "build and promote conservative public policies." More recently, news outlets have reported that a number of Democratic presidential candidates for the 2020 election are meeting with, and hiring, strategists from progressive advocacy groups.[2] In at least two of these examples, a large part of the expertise sought by the candidates is in formulating a policy platform that will appeal to voters. What is less clear is whether such advocacy ultimately benefits voters, or instead, the particular interests being represented. The results of this chapter provide some hope by showing that the two are not mutually exclusive, at least in theory.

1.2 Related Literature

In the benchmark model of cheap talk between one sender and one receiver ((Crawford and Sobel, 1982)), an important result is that messages convey information if and only if the sender's and receiver's preferences are sufficiently aligned. This carries forward to models with two or more receivers, as long as the receivers' actions enter the sender's utility independently ((Farrell and Gibbons, 1989)[3]). This model's main departure from previous ones is in introducing political competition, à la Hotelling-

[2] *The New York Times*, "2020 Democrats Import Grass-Roots Activism Into Their Campaign Staffs", March 2019.

[3] In (Farrell and Gibbons, 1989), public messages may have a disciplining effect if the total incentive to tell the truth (summed over both receivers) in each state is positive.

Downs[4], between two receiver-candidates. With political competition, information may be transmitted even though the sender is completely biased (has state-independent preferences).

In equilibrium, the sender sends private messages to one candidate only; public messages are uninformative. In the language of (Farrell and Gibbons, 1989), this is "subversion". Whereas in their paper, public messages lead the sender to prioritise one receiver over the other, here lying to both candidates is actually more tempting than lying to just one. In equilibrium, the existence of an uninfluenceable candidate is crucial because it constrains the sender as to what policies could possibly win.

Other papers have dealt with state-independent preferences in different ways. For example, (Chakraborty and Harbaugh, 2007) and (Chakraborty and Harbaugh, 2010) use multidimensional states. The former shows the sender may credibly rank issues (dimensions), while in the latter, the sender may trade off issues, by partitioning the state space such that the induced actions leave her indifferent. (Lipnowski and Ravid, 2017) allow the sender to garble information and the receiver to randomise between actions. All three papers feature one receiver, while my model has two. Furthermore, in the present chapter, the sender is not always indifferent between messages, because the true state indirectly affects her payoff. Although she always prefers a high policy to a low policy, the high policy can only win the election when the state is high enough, in which case she strictly prefers to recommend it.

To my knowledge, the only other model that combines cheap talk and elections is (Kartik and Van Weelden, 2019), where the cheap talk is done *by* candidates to reveal their ideological biases to the electorate. However, this chapter is also related to a large section of the political economy literature that tries to explain policy divergence across parties or candidates. This literature developed as a response to criticisms of the "median voter theorem" in Hotelling-Downs — a prediction inconsistent with

[4](Hotelling, 1929); (Downs, 1957)

real world observations. See, for example, (Roemer, 1994), (Besley and Coate, 1997), (Callander, 2005), (Kartik and Mcafee, 2007), (Kamada and Kojima, 2014), to name only a few. My cheap talk model predicts that, between ex-ante identical candidates, one may receive better information and diverge from the status quo policy. In particular, divergence occurs when the information suggests that voters have more extreme preferences than originally expected.

Last but not least, this chapter and the next contribute to the lobbying literature. (Grossman et al., 1994) and (Grossman and Helpman, 1996) show how a special interest can influence policy using campaign contributions. (Grossman and Helpman, 2001) provide an overview of the theoretical literature on lobbying, using pecuniary or informational means, including a chapter where they discuss cheap talk models in the context of political influence. Other papers address informational lobbying, for instance (Potters and van Winden, 1992), which models it as a signalling game, and (Austen-Smith, 1993), where lobbyists choose whether to acquire information before lobbying a committee (who sets the agenda) and the House (who vote under closed rule). This chapter contributes to the literature by studying informational lobbying at an earlier stage, namely the election process, before candidates even announce their policy positions.

1.3 Model

To model political competition, I use the Downsian model where two political parties, or candidates, each commit to a policy platform, which are then subjected to the vote of the public. The policy platform closest to the median voter's bliss point must win; in case of a tie, each wins with 0.5 probability. I assume that candidates do not know the exact location of the median voter, denoted as θ, but have some common prior over the policy space. The sender, however, knows θ exactly, and is able to send private messages to the candidates before they commit to a policy platform. Candidates are purely office-motivated, i.e. only care about winning. The sender is completely biased and wants the winning policy to be as high as possible, irrespective of θ.

There are three players: the sender, S, and two candidates (receivers), $i = 1, 2$.

$F : \Theta \to [0, 1]$ is the common prior over θ. Assume F has full support on $\Theta = [0, 1]$and continuous density f.

The timeline of the game is as follows:

1. S observes θ. Then S sends a private message m^i to each candidate.

2. Candidate $i \in \{1, 2\}$ observes m^i and chooses $a^i \in [0, 1]$. The candidate preferred by the median voter wins, and their policy platform is implemented. In case of a tie, each candidate wins with probability 0.5.

Assume that the median voter has symmetric single-peaked preferences with bliss point θ.

Define the winner of the election, w, which is a random variable drawn with equal probability from the set of potential winners, $\arg\max_{i\in\{1,2\}} -(a^i - \theta)^2$. So w is either a degenerate random variable at one of the candidates, or takes each value with probability 0.5.

Now we are ready to define payoffs. As usual in any cheap talk game, we can define payoffs based solely on the receivers' actions.

The sender's payoff is each candidate's policy weighted by their probability of winning:

$$\begin{aligned} U_S(a^1, a^2; \theta) &= E(u_S(a^w) \,|\, a^1, a^2; \theta) \\ &= Pr(w = 1|\, a^1, a^2; \theta)\;\; u_S(a^1) \;\;+\;\; Pr(w = 2|\, a^1, a^2; \theta)\;\; u_S(a^2) \end{aligned}$$

where $u_S(.)$ is a strictly increasing and continuous. In other words, S would like the winning policy to be as high as possible.

Note that, once we know a^1, a^2 and θ, each candidate's probability of winning is either degenerate or, in the case of a tie, half.

Each candidate's payoff is proportional to their probability of winning:

$$U_i(a^1, a^2; \theta) = Pr(w = i \,|\, a^1, a^2; \theta)$$

A strategy for S is a pair of message mappings $\sigma_{Si} : \Theta \to M$, for $i = 1, 2$. A strategy for each candidate i defines a policy platform for each message: $\sigma_i : M \to [0, 1]$. Therefore, on the equilibrium path, each receiver chooses a policy equal to $a^i = \sigma_i(\sigma_{Si}(\theta))$.

1.4 Equilibria

I consider pure strategy weak Perfect Bayesian Equilibria (henceforth "equilibria") of this game.

Lemma 0. Babbling is always an equilibrium. There always exists an equilibrium where both candidates choose a policy platform equal to $\theta_{\frac{1}{2}}$, the median of F, irrespective of messages. (This is a standard result in cheap talk, and holds true whether messages are public or private. The interest lies in whether any other equilibria exist).

Proposition 1. No public messages. If we restrict the sender's strategy to public messages (i.e. $\sigma_{S1} = \sigma_{S2}$), then the only equilibrium is babbling, and both candidates always choose $\theta_{\frac{1}{2}}$ regardless of messages.

Proposition 2. Babble to one, talk to the other. On the equilibrium path, at least one candidate's strategy must be independent of messages, i.e. $\exists U \in \{1,2\}$ such that $\sigma_U(\sigma_{SU}(\theta)) = a_U \quad \forall\theta$.

Proposition 3. Uninformed candidate competes at lower end. Let U denote the candidate whose strategy is independent of messages, and let R denote the other. Then $a_U = \inf A_R, \quad \text{where } A_R = \{a_R : \exists\,\theta\,\text{s.t. } a_R = \sigma_R(\sigma_{SR}(\theta))\}$.

Proposition 1 states that, if the sender can only send public messages, there can be no meaningful information revelation in equilibrium. Since messages are public, both candidates must always update their beliefs in the same manner. Hence, both candidates must converge to the median of their common posterior, as in Hotelling-Downs. Suppose there were an equilibrium with on-the-equilibrium-path messages that could induce two distinct posterior beliefs with distinct posterior medians. Then since the sender is completely biased, they would never induce the posterior with the lower median.[5] This intuition is similar to the classic cheap talk model with large bias ((Crawford and Sobel, 1982)). In essence, with public messages, it is as if we were back to a one-receiver model.

With private messages, however, some information may be transmitted. In particular, the sender sends meaningful messages to only one sender. If we compare to (Farrell and Gibbons, 1989), this is exactly what they refer to as "subversion": the sender privately reveals information to one receiver but not to the other, and if messages have to be public, no information is revealed. Note that the intuition is completely different, however. In their paper, the two receivers' actions enter the sender's utility additively — they do not interact. Here, the strategic interaction between receivers is precisely what drives this result. Under public messages, we are back to the Crawford-Sobel world, where the sender, being too biased, must babble. With private messages,

[5] The complete proof is slightly more involved. See Appendix II.

however, the presence of the uninformed candidate (who chooses an unfavourable policy) is what compels the sender to help the other candidate win.

The uninformed candidate's policy is now a lower bound on what the sender can achieve, as the sender would never recommend a lower policy platform to the other candidate. Any policy that the sender recommends must compete against the uninformed candidate's low policy, hence extreme policies can win only if the median voter prefers them to the low policy. Therefore, political competition provides endogenous credibility to the sender. If the sender recommends a policy platform that is too high above the median voter, that policy will lose the election, and the lowest policy (led by the uninformed candidate) will be implemented instead. The sender's best response therefore often involves recommending the highest (most distorted) policy platform that can still win the election.

Proposition 4. Uninformed policy no higher than the median of F. $a_U \leq \theta_{\frac{1}{2}}$.

In the absence of the sender, both candidates choose the median of F. This is the median voter theorem, where both candidates are centrists. In the presence of a biased sender, polarisation occurs, first because one candidate sometimes receives information which leads him to choose higher policies, and second, because the other candidate may also diverge from the center of the prior distribution. This is because, in equilibrium, the uninformed candidate can anticipate what messages the sender might send to the other candidate. Proposition 4 states that the uninformed candidate must, in fact, move weakly in the opposite direction, i.e. lower. In other words, the existence of the sender causes ex-ante identical candidates to diverge from the center, often in opposite directions. One candidate receives recommendations from the sender and his policies reflect the sender's bias, within the constraints of electability. The other candidate receives no additional information (save his prior) and chooses a policy platform that the sender dislikes.

One potential benefit to voters is the broader choice available to them. Instead of having to choose between two identical policy platforms, they sometimes have two different policies to choose from. In particular, the uninformed candidate's policy provides an alternative against the high policies chosen by the other. This suggests that the median voter's welfare is not necessarily worsened by the presence of the sender. Indeed, in the next section, I show that in the uniform-prior case, both sender and median voter benefit from cheap-talk lobbying.

Proposition 5. Finitely many messages. Assume $f(\theta) > 0$ for all $\theta \in (0, 1)$. The set of policy platforms for R that are induced in equilibrium, $A_R = \{a_R : \exists\,\theta \text{ s.t. } a_R = \sigma_R(\sigma_{SR}(\theta))\}$, is finite.

Proposition 5 states that, the set of messages sent in equilibrium is finite. This implies that full revelation is impossible, and in fact the extent of information revelation is quite limited, as in (Crawford and Sobel, 1982). The intuition for the proof is as follows: As we move along the state space, the sender switches to a higher policy recommendation as soon as that policy is able to win, even by a tiny margin. If the set of meaningful messages (or recommendations) is large, it means that the sender is often inducing very close elections where the median voter is almost indifferent between the uninformed candidate's low policy and the other's high policy. But then, the uninformed candidate can deviate upwards, slightly closer to the median voter, and capture all these close elections.

1.5 Uniform Prior

To see what an equilibrium of this game looks like, we must specify a prior distribution for θ. It is helpful to examine the uniform distribution, which has been extensively studied in previous models.

Lemma 1. Two messages. If $\theta \sim Uniform[0,\, 1]$, then A_R is at most a doubleton. This means that in equilibrium, the sender recommends one out of a set of at most two policy platforms to the informed candidate.

Lemma 2. Interval equilibria. If $\theta \sim Uniform[0,\, 1]$, the following is an equilibrium iff $a_\ell \in [\frac{1}{4}, \frac{1}{3}]$ and $a_h = 3a_\ell$.

Rename the candidates $R, U \in \{1, 2\}$, where $R \neq U$, to denote who receives information (R) and who receives a babbling message (U). Fix some arbitrary $m_U \in M$. Also fix a partition $\{M_\ell, M_h\}$ of M, and arbitrary elements $m_\ell \in M_\ell$ and $m_h \in M_h$.

$$\sigma_S(\theta) = \begin{cases} (m_U, m_\ell) & \text{if } \theta < \frac{a_\ell + a_h}{2}, \\ (m_U, m_h) & \text{otherwise.} \end{cases}$$

$$\sigma_U(m) = a_U = a_\ell \quad \text{for all } m \in M, \quad \text{with prior belief after every } m.$$

$$\sigma_R(m) = \begin{cases} a_\ell & \text{for all } m \in M_\ell, \quad \text{with belief } \theta \sim Uniform[0, \frac{a_\ell + a_h}{2}], \\ a_h & \text{for all } m \in M_h, \quad \text{with belief } \theta \sim Uniform[\frac{a_\ell + a_h}{2}, 1]. \end{cases}$$

Lemma 3. Sender-optimal equilibrium. If $\theta \sim Uniform[0,\, 1]$, then the following is an equilibrium, and it achieves the highest equilibrium expected payoff for the sender.

Fix some arbitrary $m_U \in M$. Also fix a partition $\{M_1,\, M_2\}$ of M, and arbitrary elements $m_1 \in M_1$ and $m_2 \in M_2$.

$$\sigma_S(\theta) = \begin{cases} (m_U, m_1) & \text{if } \theta \in [0.25, 0.75), \\ (m_U, m_2) & \text{otherwise.} \end{cases}$$

$$\sigma_U(m) = 0.5 \quad \text{for all } m \in M, \quad \text{with prior belief after every } m.$$

$$\sigma_R(m) = \begin{cases} 0.5 & \text{for all } m \in M_1, \quad \text{with belief } \theta \sim Uniform[0.25, 0.75), \\ 1 & \text{for all } m \in M_2, \quad \text{with posterior given } \theta \in [0, 0.25) \cup [0.75, 1]. \end{cases}$$

If u_S is linear, this equilibrium gives the sender expected utility $E_\theta(U_S(\boldsymbol{\sigma}^*; \theta)) = 0.625$.

Lemma 1 states that, under uniform prior, the informed candidate has at most two possible policy platforms on the equilibrium path. This means that in any non-babbling equilibrium, the information that is transmitted to that candidate w.l.o.g. takes the form of a two-cell partition of Θ. Lemma 2 characterises equilibria where the cells are intervals of the state space. Such equilibria are quite simple, and can be completely characterised by the two induced policy platforms, a_h and a_ℓ. Proposition 3 immediately implies that the uninformed candidate chooses the lower policy platform, a_ℓ. Furthermore, the cutoff between the two intervals is at the midpoint of a_h and a_ℓ. In other words, the sender recommends the higher policy platform if and only if the median voter will choose it over the lower policy platform. If θ is below that cutoff, the higher policy cannot win against the lower policy anyway, so the sender might as well recommend the latter. This is what one might call pragmatism from the part of the sender: recommending the extreme policy only when it has a chance of winning.

However, as will be shown by Lemma 3, not all equilibria feature a partition of the state space into intervals. In particular, in the equilibrium that gives the sender the highest utility, the two messages convey information about the *intensity* of voter preferences: message m_1 is sent if θ is moderate, and m_2 is sent if θ is extreme. After receiving the extreme message, the informed candidate is unsure whether voters favour extremely low or extremely high policies, so he might as well choose a very high policy (equal to 1), and win half of the time. The uninformed candidate always chooses the moderate policy ($a_U = \frac{1}{2}$), and so does the informed candidate when he receives the moderate message.

However, non-interval equilibria turn out to be fragile against small perturbations of the model. In Appendix I, I examine robustness properties when the sender does not perfectly observe θ. I find that informational imperfections worsen the sender's credibility. As a result, only interval equilibria survive.

It turns out that conveying information about the intensity — but not the direction — of voter preferences is, in general, a good strategy for the sender (when possible). This structure of messages will reappear in Chapter Two, when we discuss Bayesian persuasion. In that setting, the sender has perfect credibility, which allows them to take this strategy to the extreme. I will show that the optimal strategy for the sender is to reveal to both candidates exactly the intensity of voter preferences, i.e. how far θ is from its median. Since candidates receive no information about the direction (left or right) of θ from the median, they might as well go with the higher policy and win with probability half.

Lemma 4. Welfare If $\theta \sim Uniform[0, 1]$ and u_S is linear, then any non-babbling equilibrium gives weakly higher utility to the sender and the informed receiver, and strictly higher utility to the median voter, than the babbling equilibrium.

Lemma 4 states that the median voter's welfare is improved by the presence of the sender. The sender distorts policy, but also allows one candidate to tailor his policy platform according to voter preferences. This effectively creates more choice for the median voter, compared to babbling where candidates only ever propose one policy platform. Overall, the median voter is better off.

1.6 Why Skewness Matters

In general, the prior distribution over the median voter θ could be anything. However, the shape of that distribution matters for information revelation, because it affects candidates' incentive to deviate. Here is a stark example: if the probability density f is strictly increasing in θ, then babbling is the only equilibrium. To understand why, one must realise that the sender's credibility depends crucially on the uninformed candidate choosing a low policy. The necessity to win the election is what constrains the sender to recommend a policy platform that is not too far from the median voter. Hence, the sender and the informed candidate's incentives align — winning against the uninformed candidate is their common objective. However, for this to be an equilibrium, the uninformed candidate must not have an incentive to deviate. In particular, higher policy platforms must not be too attractive. If the prior distribution is too negatively skewed, the median voter is likely to prefer high policy platforms, and the uninformed candidate always prefers to move slightly higher to appeal to the median voter.

One interpretation is that the sender can only affect policy when their information carries unexpected "good news", i.e. when voters are more extreme (favour higher policies) than was expected under the prior. If voters were ex-ante expected to favour high policies, both candidates would already be choosing high policies even without the sender's messages. In that case, there would be little room for the sender to further increase policy.

On the other hand, if f is decreasing in θ, high policy platforms are unattractive, unless the sender provides information to the contrary. The sender is then useful (to one candidate) in flagging up opportunities to win with a high policy platform. In such cases, there always exists an equilibrium where messages are informative. The same turns out to be true if, instead, f is single-peaked and symmetric. The proof is constructive, with an equilibrium similar to Lemma 3, where the sender's message reveals whether θ is moderate (close to $\frac{1}{2}$) or extreme.

Proposition 6a. Babble if f is strictly increasing. If the prior density f is strictly increasing in θ, babbling is the only equilibrium.

Proposition 6b. Non-babbling equilibrium if f is weakly decreasing. If the prior density f is decreasing in θ, a non-babbling equilibrium exists.

Proposition 7. Symmetric single-peaked prior. Suppose the prior f is single-peaked and symmetric around $\theta = \frac{1}{2}$. Then a non-babbling equilibrium exists.

Figure 1.3 shows an example with a decreasing prior density, where the sender sends three distinct messages in equilibrium. (The faster f decreases, the more intervals can be revealed in equilibrium).

1.7 Conclusion

In this chapter, I set up a cheap talk model where an extremely biased sender, who has private knowledge of voter preferences, sends costless, non-committable messages to the two candidates in an election. I find that political competition has a disciplining effect that enables information transmission even by an extremely biased sender. Furthermore, I show that informative equilibria must satisfy a number of necessary conditions. Information transmission is sustained by equilibria where one candidate is uninfluenced by the sender's messages and always chooses a policy that the sender dislikes. Thus, the sender has an incentive to help the other candidate win by sending messages that are informative about voter preferences. This creates an endogenous alignment of incentives between the sender and one candidate.

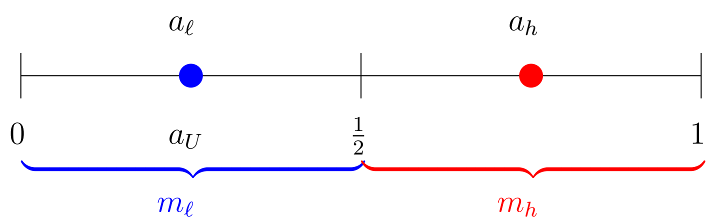
However, such equilibria may not always exist. The extent of information transmission depends on prior beliefs about the median voter's preferences. In particular, a strictly increasing prior density — indicating that the median voter is ex-ante likely to have policy preferences similar to the sender's — precludes information transmission in equilibrium. On the other hand, if the prior density is decreasing, informative equilibria exist. This may seem counter-intuitive. However, one must remember that, in any informative equilibrium, the uninformed candidate must be incentivised, based on his prior, to choose a low policy which the sender does not like. This means that the prior must assign high probability to low states.

At the exact boundary between the above two cases lies the uniform distribution, which has been extensively studied in the literature. Consequently, informative equilibria exist, but only just so. No more than two policy platforms are chosen on the equilibrium path, which means that the information transmitted is extremely coarse. A natural interpretation is that the sender makes one of two recommendations ("low policy", "high policy") to one candidate, while the other candidate always chooses the lower policy. I show that the presence of the biased sender in this setting improves the median voter's welfare.

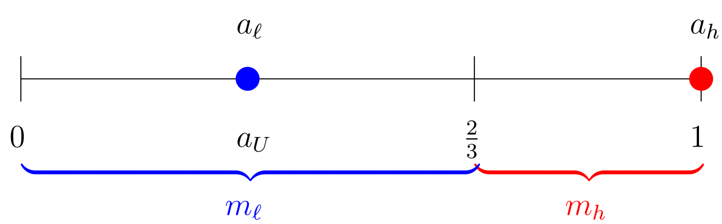
These predictions match real-world observations in three important ways. First, the model explain why candidates would trust advisers that are known to have extreme policy motives — a prediction that standard cheap talk models are unable to deliver. Second, it provides intuition for why advisers (especially biased advisers) often exclusively serve one candidate: an adviser who influence both candidates via cheap talk cannot be trusted, as she would lie. Third, this leads naturally to the conclusion that an adviser that influences one candidate but not the other would lead to policy differences between candidates, due in particular to the asymmetric information they receive. In some cases, such as in Section 1.5, this policy divergence strictly benefits the median voter.

Finally, a key takeaway from this chapter is that political institutions dramatically shape the influence of special interests, and biased entities in general, on policy. The presence of robust, competitive elections allows for information to be extracted even from extremely biased sources, provided the latter have limited commitment power. In contrast, standard cheap talk models that ignore electoral incentives more aptly describe a benevolent dictatorship, with the well-known result that only sufficiently unbiased sources can reveal any information. Therefore, the institutional context has important ramifications on the role of special interests in politics and their effect on voter welfare.

1.8 Figures



If $a_\ell = \frac{1}{4}, \quad a_h = \frac{1}{2}$.



If $a_\ell = \frac{1}{3}, \quad a_h = 1$.

Figure 1·1: Equilibria where messages denote intervals of the state space. (Uniform prior).

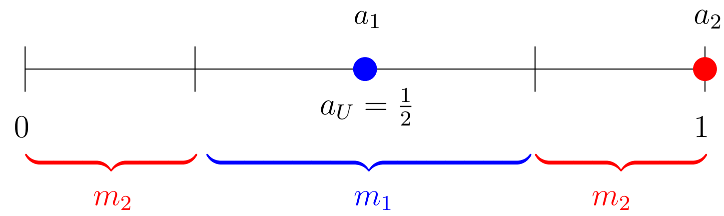


Figure 1·2: Sender-optimal equilibrium. (Uniform prior).

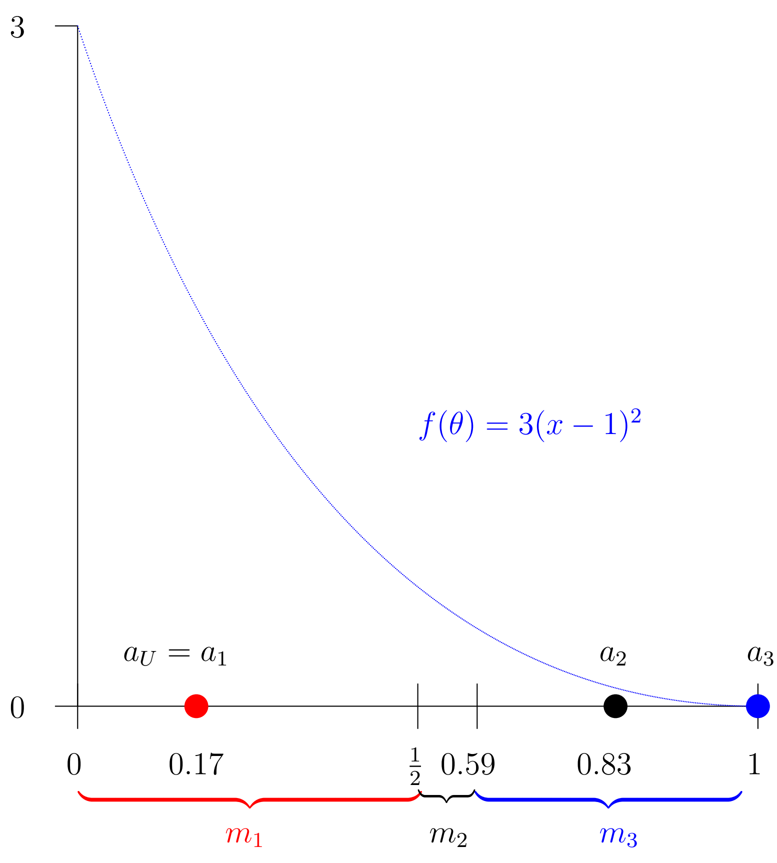


Figure 1·3: Equilibrium with decreasing f.

Chapter 2

Biased Campaign Advice: An Information Design Problem

2.1 Introduction

The previous chapter studies informational lobbying in a cheap talk context, where the sender cannot commit to the messages she will send. There, credibility — overcoming the temptation to lie — is crucial for information revelation. In the real world, however, a well-established interest group may have considerable commitment power. The present chapter studies the problem faced by an extremely biased sender who fully commits ex-ante to the information that she will send to candidates before they make their policy decision. The existing literature refers to such strategic communication under full commitment as Bayesian persuasion or information design.[1]

Commitment power may arise from strong reputational incentives (not modelled here) or from disclosure obligations, such as in the seminal (Kamenica and Gentzkow, 2011) example where a prosecutor can decide what evidence to look for but is legally bound to reveal the outcome of any investigation. In another interpretation, the sender commissions an experiment, the design of which is publicly observed, then decides whether or not to publish the resulting report. (Gentzkow and Kamenica, 2017) show that, in such situations, the report is always published, therefore yielding predictions that are identical to those of a Bayesian persuasion model.

[1] Typically, Bayesian persuasion refers to situations with one receiver whereas information design involves multiple receivers.

In games with multiple receivers, it is often helpful to imagine the sender as an "information designer" separate from the game, who is free to manipulate players' beliefs about a payoff-relevant state of the world before the game begins. The information designer is restricted solely by a Bayes plausibility condition: each player's posterior beliefs must be, on average, equal to the common prior belief on the state. It is easy to see that information design must allow the sender to achieve weakly higher utility than under cheap talk, since an equilibrium under cheap talk can also be replicated by an information designer. Comparing the two frameworks enables us to understand the value of commitment — how much better off is the sender under full commitment (information design) as opposed to no commitment (cheap talk).

In this chapter, I study a model similar to that in Chapter One, with the exception that the sender now has full ex-ante commitment on her message strategy. I show that commitment power drastically affects the model's predictions. In particular, whereas only private communication is possible under cheap talk, the opposite is true when the sender has full commitment. The sender's optimal strategy is shown, using numerical optimisation, to involve only public messages. Furthermore, whereas under cheap talk only coarse information can be revealed, the sender with commitment is shown to optimally reveal very precise information. This implies that the sender's utility in the information design framework is generically strictly higher than under cheap talk.

The sender's optimal strategy involves revealing to both candidates exactly how extreme voter preferences are, i.e. the quantile distance from the prior median. Hence, a message pools together exactly two possible realisations of the state of tne world: one to the left and another to the right of the median state. However, candidates are unable to distinguish which of these two policies correspond to the median voter's true preference, thus there is an equilibrium where they both choose the higher policy favoured by the sender. Since both candidates choose the same policy, the latter is guaranteed to be implemented.

It may at first not be obvious what the sender's messages would look like in real-world situations, but they could involve, for instance, the relative chances of moderate candidates as opposed to extreme ones, but in very precise terms. An example: "Only 5% of the time is the median voter's preference this far in the tails". A more intuitive, if less precise, statement might be: "The median voter wants a Bernie Sanders or a Donald Trump, nothing in between." Statements of a similar vein often appear as newspaper headlines around election time: "Moderates Can't Win the White House" (The Atlantic, February 2020), "Everything in Moderation: Why a Left-wing Nominee Would Hurt Democrats" (The Economist, November 2019), so the notion of measuring voters' divergence from the centre is not an uncommon one.

On the equilibrium path, both candidates choose the same high policy, so the implemented policy is never lower than the median of the prior distribution. This is the best possible scenario for the sender, who would like policy to be as high as possible, but not for a risk averse median voter who prefers policies close to the state of the world. Indeed, given uncertainty between two values of the state, the best policy from the median voter's perspective is a compromise between the two values, not one of the extremities. It follows that the median voter is strictly worse off under the sender's optimal strategy than they would be if the implemented policy was always equal to the prior median. In other words, the presence of the sender, although she provides information to candidates, leaves the median voter strictly worse off than under no information.

The rest of the chapter is as follows. First, I set up the sender's problem, which involves the same elements as Chapter One, except that the sender has full ex-ante commitment on her message strategy. Then, I examine a restricted version of the sender's problem by imposing that only public messages are available to the sender. Thereafter, I show by numerical optimisation that public messages are indeed optimal in the unrestricted problem. Lastly, I derive the above welfare result and discuss policy implications.

2.2 Related Literature

(Kamenica and Gentzkow, 2011) study situations where a sender designs an experiment, or equivalently, commits to a state-contingent message strategy, in order to influence the actions of a receiver. More generally, we can imagine a sender, or information designer, seeking to influence players' behaviour in a given game, by sending them signals about a payoff-relevant state. (Kamenica, 2019) and (Bergemann and Morris, 2019) provide useful surveys of the literature.

Papers featuring multiple receivers include (Heese and Lauermann, 2019) and (Gitmez and Molavi, 2018), where the receivers are voters. This chapter presents an information design problem where the two receivers are candidates in a Downsian election. Constraining signals to be publicly observed, I completely characterise the optimal strategy for the sender, and show the sender always gains from persuasion. Note that the state space in this model is uncountable (the unit interval), which means (Kamenica and Gentzkow, 2011)'s concavification tools do not apply. (Gentzkow and Kamenica, 2016) explore infinite state spaces, but assume that the receiver's action only depends on his posterior mean. In contrast, my model can be rewritten as a one-receiver problem where the receiver always selects the *median* of their posterior belief.

I show that this problem is a relaxed form of an optimal gerrymandering problem, with similarities to (Friedman and Holden, 2008). Accordingly, the optimal sender strategy has a very similar structure, pooling very high and very low states. Another paper that highlights a connection between persuasion and a seemingly unrelated problem is (Kolotilin and Zapechelnyuk, 2019), which shows that persuasion and delegation problems are often equivalent.

2.3 Model

The sender designs a message strategy, or signal structure, that specifies a joint distribution over messages (signals) contingent on the state of the world. After receiving a message, each candidate updates their belief about θ and chooses a policy platform. A candidate's strategy must be optimal given their belief about θ and the other candidate's strategy. Assuming that the sender can choose the most favourable equilibrium, we can write this as a maximisation problem subject to obedience constraints.

The sender chooses a message strategy, which specifies a joint distribution $p(.|\theta) : M \times M \to [0, 1]$ for $\theta \in \Theta$, and a strategy profile for candidates, $\{\sigma_i\}_{i=1,2}$, to maximise

$$E(u_S(a^w))$$

where a^w is the winning policy defined as previously, subject to the following obedience constraints:

For each i, for all m_i that are sent with non-zero probability for some θ,

$$\sigma_i(m_i) \in \arg\max_{a_i} \int\limits_{a_j} \int\limits_{\theta} Pr(i \text{ wins } |a_i, a_j, \theta)\ p(m_i, \sigma^{-1}(a_j)|\,\theta)\ \mathrm{d}F_m(\theta)\,\mathrm{d}a_j$$

The term $p(m_i, \sigma^{-1}(a_j)|\,\theta)\,\mathrm{d}F_m(\theta)$ in the above statement gives the joint distribution of θ, and a_j given the message m_i that player i observes. Thus, player i updates his belief not only of the state of the world, but also of what information the other candidate might have received. The statement above requires i's policy to be optimal given this updated joint belief.

2.4 Public messages with full commitment

The sender designs a signal structure, and the signal realisation is publicly observed. In other words, simply add the following constraint to the above problem: if $m_i \neq m_j$, then $p(m_i, m_j|\theta) = 0$. Since the candidates have a common prior and observe the same signal, they must therefore update their beliefs in the same manner, and each choose a median of the posterior. The best equilibrium for the sender is where both candidates converge to the highest median for each posterior.

The sender's problem can be rewritten thus:

Choose $p(.|\theta) : M \to [0, 1]$ to maximise

$$E_m(u_S(\max\ \theta_{\frac{1}{2}}(m)))$$

where $\theta_{\frac{1}{2}}(m) \equiv \{t \in \Theta :\ Pr(\theta \leq t\,|m) \geq \frac{1}{2} \text{ and } Pr(\theta \geq t\,|m) \geq \frac{1}{2}\}$ is the set of medians[2] of the posterior belief given message m.

First, I will argue that this is equivalent to a relaxed gerrymandering problem. In a standard gerrymandering problem, a principal decides how to allocate a given population with a given preference distribution $\mathcal{F} : \theta \to [0, 1]$ into constituencies of a given size. Each constituency then elects a representative, and the principal wants to maximise the average ideology of the representatives. [3] If we think of

[2]There are two ways of dealing with non-unique posterior medians. One is to only consider the highest (most favourable) median, i.e. $\hat{\theta}_m(s) = \sup \theta_m(s)$, or the lowest (least favourable). Another is to restrict the signal structure so that only posteriors with unique median are allowed. The signal structure described later in Proposition 8 solves the sender's problem in the most favourable case. However, it is easy to construct arbitrarily similar signal structures that induce unique posterior medians and achieve almost the same utility.

[3]The simplest form of this is when preferences are binary, i.e. $\theta \in \{0, 1\}$, where 0 represents a Democrat and 1 a Republican voter. Then the principal wants to optimally allocate voters so as to elect as many Republican representatives as possible. The optimal solution famously involves "packing and cracking", whereby some constituencies are overwhelmingly Democrat and some have just enough Republicans to make a majority.

our prior distribution $F(\theta)$ as being analogous to the voter population $\mathcal{F}(\theta)$ in the gerrymandering problem, and fix some finite support for $p(.|\theta)$, then the sender's choice of $p(m|\theta)$ is equivalent to allocating voters of a given preference θ to each constituency $m \in M$. In each constituency m, the elected representative is the median $\theta_{\frac{1}{2}}(m)$. The main difference here is that the support of p, instead of being a fixed number of constituencies of equal size, can be chosen to be any (potentially uncountable) set of messages. Our maximisation problem therefore has more degrees of freedom than the usual gerrymandering problem.

This connection with the gerrymandering literature is, however, helpful in formulating the optimal signal structure, which is the limiting case of the solution found by (Friedman and Holden, 2008).

Proposition 8. Sender's optimum. The following signal structure is the solution to the sender's public persuasion problem, and achieves the maximal utility equal to $E(u_S(\theta) \mid \theta > \theta_{\frac{1}{2}})$.

$$p(m|\theta) = 1 \quad \text{iff} \quad \theta = F^{-1}(\frac{1}{2} \pm m)$$

$$\sigma_i(m) = F^{-1}(\frac{1}{2} + m) \qquad \text{for all } i, m$$

In other words, the sender's optimal message strategy perfectly reveals θ's quantile distance from the median of F. The sender optimally reveals exactly how extreme voter preferences are, but not in which direction. Faced with this uncertainty, candidates might as well choose the higher value out of the two possibilities.

Under this optimal strategy, the sender achieves a utility equal to $E(u_S(\theta) \mid \theta > \theta_{\frac{1}{2}})$. Cut the prior distribution in half at the median, then throw away the lower half. Take the expectation over the remaining top half, and that is the sender's utility. Note that it is *always* strictly better than babbling, thus the sender always benefits from persuasion. This is in contrast with cheap talk, where information revelation is impossible under some priors, and with the typical persuasion problem, where gains from persuasion are not guaranteed.

Proposition 8 presents an interesting implication: the entire lower half of the prior is irrelevant to the policy outcomes induced by the sender. This should not be surprising, given the parallel to gerrymandering, since a gerrymander can always suppress up to half of the electorate by cleverly placing them in constituencies where they would be in the minority. Moreover, it allows us to draw the following inference:

Suppose that, before the Bayesian persuasion game is played, the sender is able to engage in a public opinion campaign with the aim of influencing preferences. In other words, the sender chooses a (potentially costly) action which changes the distribution of the median voter, F, in the Bayesian Persuasion game. Then what types of changes would be targeted by the sender? Any action that preserves the upper half of F is pointless. In fact, the sender should target realisations of θ above the median. In the gerrymandering case, this can be interpreted as "preaching to the choir", i.e. targeting people whose preferences are already favourable to the sender, and making them more extreme. In our case, it means inducing a thick upper tail in the distribution of the median voter.

Suppose further that the sender can only induce a symmetric distribution F. Then the more polarised F is, the better, as it entails higher $E(u_S(\theta)\,|\,\theta > \theta_{\frac{1}{2}})$. Both the lower and upper tails are made thicker, but the lower tail does not matter anyway, as it will be optimally suppressed in the Bayesian persuasion game.

2.5 Private messages with full commitment

The sender's problem becomes significantly more complex once private messages are allowed. However, if discretised, it can be solved by linear programming, for a given prior distribution. Solving numerically for the optimal sender strategy, I find that public messages are optimal even when unrestricted private messages are allowed. A sample code in R can be found in Appendix A.3.

Figures 2.4 to 2.6 show graphical renditions of the joint distribution of policies induced by the sender at optimum. First notice that only points on the 45 degree line are assigned non-zero probabilities. Furthermore, the densities reported mimic the upper tail of the prior distribution in each case (except at the exact median where discretisation leads to a slightly smaller density). This confirms that the optimal private message structure is indeed the optimal strategy described in Proposition 8.

2.6 Welfare Analysis

In this section, I study welfare implications for the median voter under quadratic utility. This is equivalent to average voter welfare if all voters have quadratic preferences $U_j = -(\text{winning policy} - \theta_j)^2$, with θ_j symmetrically distributed around θ.

While the presence of the special interest increases the amount of information that is potentially available to candidates, it may also lead to policy distortions. This is the classic bias-variance tradeoff. One could think of the implemented policy as an estimator for θ, so that the median voter's welfare is equal to the mean squared error of that estimator. It is well-known that the quadratic-loss function or MSE places equal weight on the bias and variance of an estimator.[4] Hence, a special interest is beneficial if and only if it increases policy precision around the median voter more than it increases bias.

Proposition 10 states, under public persuasion, this is never the case. In other words, with full commitment power, the sender is able to distort policy so severely

[4] This particular case is, however, slightly different from the classical framework where the parameter of interest, θ, is fixed, and variations in the estimator arise purely from sampling. Here, θ is drawn from a distribution, and the variance term is actually the covariance between θ and the policy. More precisely, we can rewrite $E(a-\theta)^2 = Var(a-\theta) + (E(a-\theta))^2$, where the first term includes the covariance, and the second term is the bias. The covariance simply measures how well a tracks θ on average.

that it outweighs any informational benefit.

Proposition 10. The median voter's welfare under public persuasion is strictly worse than under babbling.

For every message, the posterior is degenerate at two values, with the prior median between them. Under babbling, the policy is equal to the prior median, whereas under persuasion, the policy is at one extreme (the higher value). At every posterior, the average distance between the policy and θ is the same in both situations, but the variance is larger under persuasion. Since the median voter is risk averse, their welfare decreases.

2.7 Conclusion

This chapter presents an information design problem faced by a sender who can manipulate candidates' beliefs before a Downsian election. By pointing out similarities between this problem and gerrymandering, I fully characterise the optimal strategy and show that the sender will optimally reveal to both candidates precise information about the intensity, but not the direction, of the median voter's preference. This optimal strategy is shown to be strictly welfare-decreasing for the median voter. From a policy standpoint, this suggests that disclosure rules requiring all special interest communications with candidates to be made public are likely to be ineffective at improving the median voter's welfare.

This chapter stands in stark contrast with the previous chapter, suggesting that understanding special interests' degree of commitment power is crucial for making predictions. It also suggests a simple rule of thumb to distinguish between special interests with low and high commitment power: special interests that align themselves exclusively with one party or candidate are likely to have limited commitment, whereas those that interact equally with all parties are likely to have a relatively large commitment power. Furthermore, while the former may benefit the median voter, the latter always leaves the median voter strictly worse off.

The comparison between the first two chapters raises interesting points about real world special interests and their influence on American politics. For instance, an oft-discussed example is Stephen Bannon, who led Donald Trump's presidential campaign in 2016. This seems to broadly fit the description in Chapter One — a first-time campaign adviser with biased policy preferences, presumably with no commitment to truth-telling, who advises one candidate in opposition to the other. The welfare result in the previous chapter suggests that, insofar as providing information to the candidate, Mr. Bannon's involvement may theoretically have benefitted the median voter. On the other hand, an example that seems to fit Chapter Two is the National Rifle Association (NRA), an interest group representing the interests of gun owners. The NRA is a well-established institution with long-standing ties to politicians on both sides of the aisle. The predictions in Chapter Two suggest that such organisations may be detrimental to the median voter, as they induce both candidates to adopt biased policies — leaving voters with no choice. Interestingly, this line of argument has long been adopted by opponents of the NRA, who claim that, should voters be given a choice, they would opt for a greater degree of gun control. A recent rift between the NRA and the Democratic party is likely to put this argument to the test.

This leaves open an interesting avenue for future research: studying how relationships between candidates and special interests develop over time. Examining these dynamic relationships is important in understanding not only the factors that affect the longevity of relationships between politicians and their advisers, but also the informational advantages or disadvantages faced by incumbents — and thereby the persistence of political power.

2.8 Figures

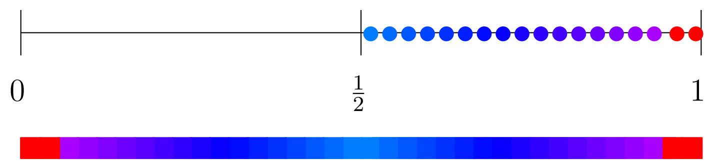


Figure 2·1: Optimal signal structure under $\theta \sim Uniform[0, 1]$. (Each square colour denotes a signal. Candidates' policy is shown by circle of corresponding colour).

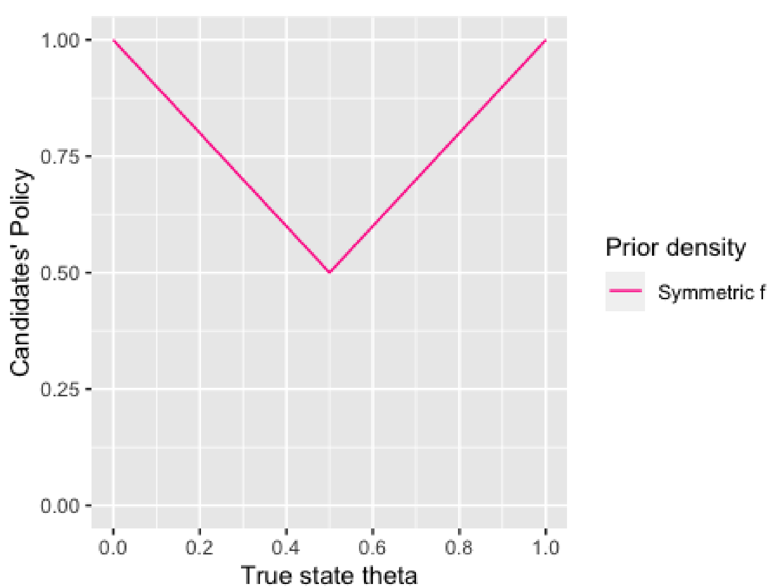


Figure 2·2: Policy as a function of median voter preferences (Symmetric prior)

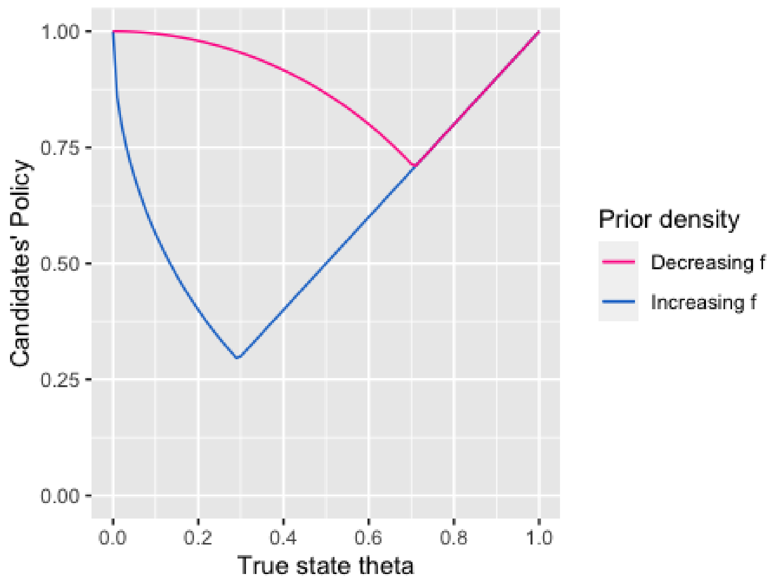


Figure 2·3: Policy as a function of median voter preferences (Skewed priors)

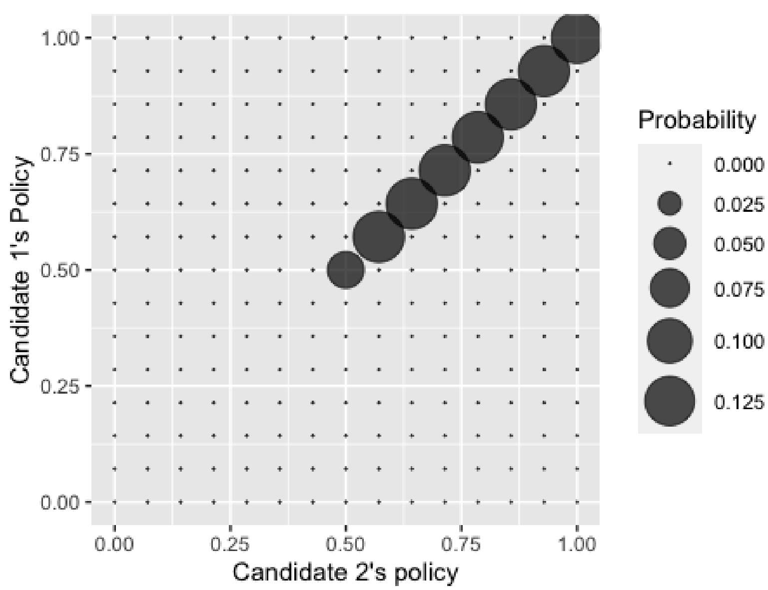


Figure 2·4: Optimal private signal structure (Uniform prior)

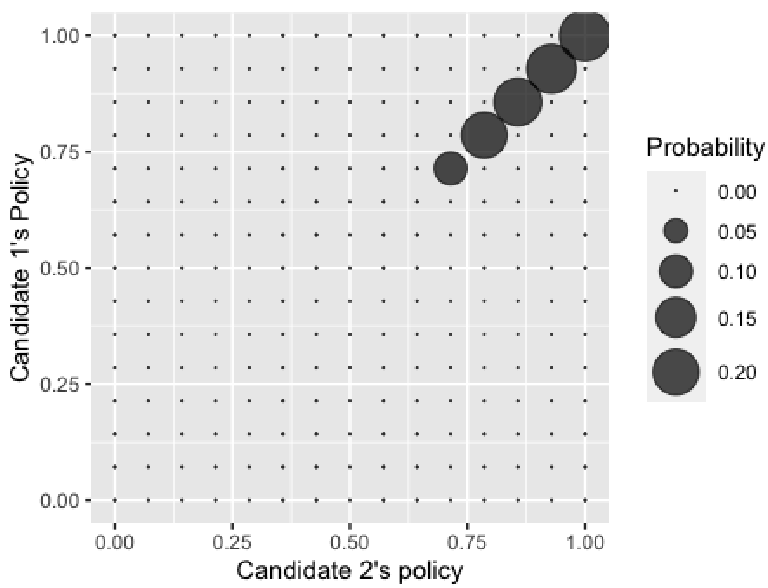


Figure 2·5: Optimal private signal structure (Prior with increasing density)

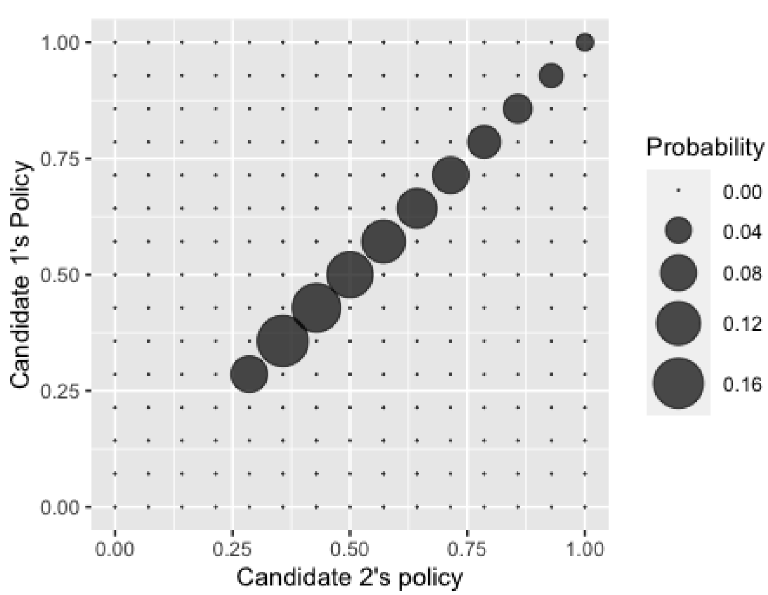


Figure 2·6: Optimal private signal structure (Prior with increasing density)

Chapter 3

Gender and Group Decision-Making: Evidence from City Councils in the United States (with Jesse Bruhn and Anna Weber)

3.1 Introduction

The participation of women in local government in the United States has increased significantly in the last decades. For instance, (Ferreira and Gyourko, 2014) report that around a third of mayoral elections attracted female candidates in the period 1995-2005, up from less than 10% in the 1970s. This has been a gradual increase rather than a result of any explicit policies such as quotas, in contrast to countries such as India. Although these numbers still reflect the continuing underrepresentation of women across the board in politics, they raise important questions about the effect on public policy and the allocation of public spending, with United States local governments spending upwards of $1.6 trillion in 2016.[1]

Economic theory offers divergent predictions about the impact of female policymakers. On one hand, the median voter theorem (Downs, 1957) states that office-motivated politicians of either gender will, in equilibrium, enact policies that conform to the

[1]Source: U.S. Census Bureau. The reported number is for local general expenditures, which include spending on schools, health care services, and general administration (among other activities in the general government sector) but exclude government-run liquor stores, utilities, and insurance trusts.

median voter's preference. Hence, no change is to be expected from an increase in female representation. On the other hand, more recent papers, such as (Besley and Coate, 1997)'s "citizen candidate" model, recognise that policymakers have preferences of their own and may not be able to commit to not implementing them once elected. This implies that, if male and female citizens have differing policy preferences, they may govern differently.

A large body of literature documents precisely such an effect in developing countries. Researchers working in this area have consistently found that female representation not only shifts the composition of public spending towards categories associated with broad female support, but can also have downstream effects on the aspirations and educational attainment of school age girls (See (Chattopadhyay and Duflo, 2004),(Beaman et al., 2009), and (Beaman et al., 2012)). However, studies that have looked for similar effects in western democracies have by and large come up empty handed.(Ferreira and Gyourko, 2014) find that in the United States, when a woman wins a close mayoral election against a man, there is no impact on the composition of public spending.

This chapter is the prelude to a larger project examining the effects of female representation in city councils in the United States. The municipal setting provides a novel dimension of female representation — intensive rather than extensive — in that an additional female council member only marginally affects the gender composition of the policymaking group. This is in contrast to most existing empirical studies on gender. Furthermore, while presently we focus on the effects on policy outcomes such as the size and composition of municipal spending, our aim is more broadly set on understanding not only the outcomes but the mechanisms that lead to them.

This chapter has three main contributions. First, we construct a new dataset linking city council election information with data from the records of proceedings of council meetings. Second, by exploiting close elections between male and female candidates, we estimate the effect of gender on policy in a new context — the city council — and find that an exogenous increase in the number of female council members has a weakly significant negative effect on per capita expenditure, but no significant effect on the composition of municipal spending. Third, we find no significant effect on the participation or success of other women in future elections.

Our results mirror existing literature in a few interesting ways. The null effect on spending composition recalls (Ferreira and Gyourko, 2014)'s findings on the effect of female mayors, with a number of possible explanations. For example, men and women may not differ significantly in their policy preferences. Available evidence, however, contradicts this. The American National Election Studies reports some differences in male and female preferences for local goods, with results suggesting that women care relatively more about police, local elections and welfare programmes. Another potential explanation may be that the median voter theorem applies in this western setting, unlike in other developing countries that have been studied. A crucial element of the Downsian model rests in politicians' ability to commit to a policy that is independent of their own preference. If western democratic institutions allow greater commitment power to politicians — for instance, through media scrutiny and greater accountability — the median voter theorem is more likely to hold. This, however, fails to explain the following result.

We find some evidence that increasing the number of female councillors decreases per capita total spending. Although the statistical significance of this effect is somewhat sensitive to bandwidth choice, the estimates are stable and economically large. This holds even after controlling for candidates' previous experience in running for and holding office. Our result is closely related to a strand of literature that examines, both theoretically and empirically, the relationship between diversity and public good provision, and more broadly between diversity and conflict (see, for example, (Alesina and La Ferrara, 2005) and (Ray and Esteban, 2017) for useful reviews). (Alesina et al., 1999) present a hypothesis that increased diversity in a society leads to lower public good provision through heightened disagreement about how to allocate public funds. In the context of city councils in the United States, (Beach and Jones, 2017) find that increased ethnic diversity in city councils in California decreases total municipal spending on public goods, which they interpret as evidence of a gridlock within the council.

In the following section, we develop a model of legislative deliberation which fleshes out two pieces of intuition on female representation (or more broadly, minority representation). The first, which we refer to as "self-censoring", relates to the disincentive for a member of a minority to propose unpopular policies or voice her opinions. There is some experimental evidence on self-censoring; for example, (Coffman, 2014) finds that even when individuals are knowledgeable in a specific area, they systematically under-contribute to a group decision making process when the subject is outside of their stereotypical gender domain. A second effect, which we refer to as "pushback", pertains to the fact that the majority may react strategically to a change in group composition. This may lead to increased disagreement within the group, with potentially negative consequences. For instance, (Bagues and Esteve-Volart, 2010) document, in the context of Spanish civil service positions, that women who are randomly assigned to an evaluation committee with an above average share of women are, in fact, less likely to be hired. The main way in which this manifests in our model is with a decrease in men's demand for public spending.

Whereas existing studies, including the present analysis, focus on policy outcomes, such as municipal spending budget and composition, our broader aim is to provide a glimpse into the decision-making process, in order to detect potential changes in strategic behavior and disagreement. Therefore, as we walk through this chapter, it is helpful to keep in mind the next part of the project, which deals with records of proceedings of city council meetings. By examining the minutes of council meetings, we can measure the effect of female representation on the frequency with which members of different genders participate in the discussion, how often they propose or second motions, voting patterns and meeting length. This will enable us to get a more complete understanding of gender dynamics in group policymaking.

3.2 Theoretical Framework

The following is a stylised model that illustrates our intuition for "self-censoring" and "pushback". Suppose that city council budgets are determined in a two-stage deliberative process: first setting a total budget (either *high* or *low*), then choosing between three possible allocations over public goods (m, f or a compromise good x). The first stage is a simple majority vote between the two budget levels, and the second involves a proposal followed by voting.

There are N members with stochastic preferences over allocations, which have two possible realisations:

$$m \succ x \succ f \qquad \text{or} \qquad f \succ x \succ m$$

For simplicity, let each member's preferences be independently distributed, and identically within each gender. Assume that men are strictly more likely to have preferences $m \succ x \succ f$ than $f \succ x \succ m$, and the opposite for women.

Suppose that after the budget is decided, members' preferences over allocations are realised and a member is randomly chosen by nature to make a proposal. The chosen member can either make a proposal at some marginal cost $c > 0$ or decline to make a proposal. Let the compromise good be the status quo, which is implemented if no proposal is made or if a proposal is rejected.

Then each member i will make a proposal reflecting his or her true preference $\theta_i = \arg\max_{a\in\{m,f\}} u_i(a)$ if and only if

$$Pr(\,\theta_i \text{ passes the vote}\,)(u_i(\theta_i) - u_i(x)) - c \geq 0$$

and decline to make a proposal otherwise.

Examining the expression above leads us to our first prediction. If men constitute the majority in the council, then keeping everything else constant, a member who favours the allocation m anticipates a higher probability of success for his or her proposal. Therefore, men are overall more likely to make proposals than women.

Prediction 1. A male member of the council proposes motions more frequently than a comparable female member.

Furthermore, as the gender composition of the council shifts to include more women, the left-hand side in the above expression becomes larger for members who favour allocation f, and smaller for those who favour allocation m. Two things may happen: either it becomes worthwhile for members of all preferences to make proposals, or no one makes a proposal. The former is a situation where diversity opens up the conversation for all, whereas the latter shows increased polarisation may shut down any discussion in favour of a compromise allocation.

In either case, the shift in gender composition leads to a movement away from the male-oriented good m. From an ex-ante point of view, male members who control the majority anticipate a lower expected utility. This leads to the following prediction.

Prediction 2. Increased female representation leads to a weakly lower total budget, and a weak decrease in the probability that allocation m is chosen. The probabilities with which allocations f and x are chosen may increase or decrease, depending on whether female representation encourages or dampens members' incentive to make proposals.

In subsequent sections, we describe the institutional setting for our empirical study, followed by our identification strategy, and present our main results.

3.3 Setting

We focus on local governments at the municipal level: city/town councils, village boards of trustees, and boards of alderpersons. Often the format is specified by the state constitution. The jurisdiction and responsibilities of councils vary, but may include police/fire protection, sanitation, local roads, parks and recreation, libraries, and local zoning. (However, they do not include the funding or operation of schools, which fall under the jurisdiction of school districts).

A typical council in our data consists of 4-7 members and a "mayor". About half of the council is elected every 1-2 years, either at-large or by district. Council members are usually elected on non-partisan basis. We include the mayor as a member of the council where available, but exclude elected administrative positions like Secretary, Clerk, etc.

3.3.1 City Council Election Data

We collected election data by sending Freedom of Information Act requests to 3,513 cities. The list of cities was taken from the Annual Survey of State and Local Government Finances (from the US Census Bureau), and consists of all municipalities included in the 2009-2013 subsample. Individual FOIA requests were sent to local governments to collect election data from 2008 to mid- 2017. Responses were received in a variety of formats, including pdfs and links to websites. So far, we have entered election data for 223 cities. For each city and election, we have the election date, position (mayor/council), first and last names of candidates, vote counts and elected candidates.

We supplement our data with the California Election Data Archive (CEDA), also used in (Beach and Jones, 2017), which provides the names and number of votes for every candidate in every local government election occurring in our period of interest.

We requested candidate gender in our FOIA request but it was only provided by a small number of cities. In order to impute a gender to each candidate, we follow a name-based procedure inspired by Ferreira and Gyourko (2014). For each first name, we compute the fraction of individuals born in 1950-1980 with that name who are female, using name counts by year of birth and gender from the Social Security Administration. For the worst winner and best loser in contested elections, we use a 95% threshold to assign gender; if ambiguous ($>$ 5% and $<$ 95% female), the individual's gender is ascertained by an internet search. For rest of council, we use a 50% threshold to assign the gender.

For most of the sample, we are able to construct the council composition from regular election data. (This excludes special elections and appointments between regular elections after resignations or deaths). The modal number of elections to turn over the whole council is 2. Therefore, starting from the second election included in our data, we can reconstruct the council from winners of the past 2 elections.

Table 3.1 shows summary statistics related to the elections in our sample. Around 80% of cities have had at least one race where the worst winner and best loser are of opposite genders, and such races account for around 22% of all races. This leaves us with approximately 1,400 gendered races. Figure 3.1 is a histogram of the vote margins in the closest gendered race for each city, showing that a large number of cities have had close gendered elections in our dataset.

3.3.2 Outcome Variables

We analyze municipal spending data from the Annual Survey of State and Local Government Finances from the US Census Bureau, which provides data on expenditures and revenues by category. This we supplement with data from annual city budgets for Californian cities, made available by the California City Controller's Office. We also construct some outcome variables using the election data described above.

Table 3.2 shows summary statistics for expenditure in our sample in 2012, which is the midpoint of our period of interest. Californian cities and the cities in our sample show on average much higher expenditure than the typical city in the United States census, so our results may not be representative of smaller or less affluent cities in the United States. Cities with gendered elections also have higher expenditure on average compared to a typical city in our sample.

3.4 Empirical Approach

We would like to estimate the causal effect of a change in female representation on a variety of outcome variables. Absent any endogeneity concerns, the simplest possible specification is the following panel OLS:

$$Y_{ct} = \beta_1 treat_{ct} + \beta_2 members_{ct} + \alpha_c + \tau_t + u_{ct}, \tag{3.1}$$

where subscripts c and t refer to the city and year of observation, and $treat$ could be an indicator for any woman on the council, the count of female members, female percentage of the council, etc. For a binary treatment, this is essentially a difference-in-difference estimator.

However, the estimated coefficient β_1 is likely to be biased, since we expect that changes in female representation are correlated with changes in political views within a city, or with council collaborative culture.

To deal with this concern, we implement a regression discontinuity approach, focusing on races where the gender of the marginal council member is determined by the result of a close race. Since an election may include multiple races and each race may have multiple winners, the analysis is conducted at the race level and considers only races where the worst winner and best loser are of opposite genders. For each race, we calculate the female margin of victory as a (signed) percentage such that a margin of zero indicates a tie. This allows us to zoom in on close gendered races where, around the cutoff of zero, the gender of the marginal council member is approximately random.

For observations where the margin of victory is within a bandwidth around a vote margin of 0, estimate:

$$Y_{rct} = \beta_1 femwin_{rct} + \beta_2 margin_{rct} + \beta_3 (femwin_{rct} * margin_{rct}) + \alpha_c + \tau_t + u_{rct} \quad (3.2)$$

where $femwin_{rct}$, the treatment variable, is an indicator for whether the female candidate wins the race, whereas $margin_{rct}$ is the running variable.

Figure A.2 shows that the gender composition of the council is discontinuous in the vote margin. This indicates that the regression discontinuity is indeed valid. Furthermore, Tables A.12 and A.13 report the results of continuity checks for a number of predetermined and other variables, including election parameters and city population. Figures A.3 to A.7 show the corresponding binned scatter plots. Testing for vote manipulation around the cutoff (in accordance with Cattaneo, Jansson and Ma (2019)) returns a test statistic of 1.5341, with a p-value of 0.1250. The corresponding density plot is shown in Figure A.8.

3.5 Results

Table A.1 and A.2 show the results from equation 3.1 where the outcome variables are measures of total and per capita municipal expenditure. These tables show correlations between total expenditure and various measure of female representation. In particular, real per capita expenditure is strongly negatively correlated with the fraction of women council members, with each percentage point increase associated with $72 higher per capita expenditure. Table A.3 shows the outcomes of similar regressions on the shares of expenditure devoted to different categories of spending. There is a strong negative correlation between a number of measures of female representation and the shares of expenditure allocated to health and hospitals, and roads and parking. However, these estimates can in no way be interpreted as causal. There are strong arguments for female representation being endogenous to expenditure allocations. Although we include city fixed effects, changes in female representation across time within a city is likely to be highly correlated with other changes in the political climate, leading to correlations with expenditure patterns.

Therefore, we turn to equation 3.2 and report RD estimates showing the causal impact of a female win in a close gendered race on expenditure. We report the results estimated at the optimal bandwidth as per (Calonico et al., 2014), as well as half and twice the optimal bandwidth. Table A.4 shows some evidence that a female win decreases per capita expenditure by around 9% to 11%. The statistical significance of this estimate varies by specification, but the estimated effect is consistently above 9% across the board. Our estimates are of very similar magnitude to the ones found by (Beach and Jones, 2017), who report a 10% decrease in public goods expenditure per capita following a win by an ethnic minority candidate. Table A.5 reports the RD estimates under the optimal bandwidth under polynomial specifications. Figures 3.1 and 3.2 are binned scatter plots of the same expenditure variables or residuals after taking out state and year fixed effects.

One potential concern for the validity of the RD estimates lies in the fact that women are often less experienced in holding office than men. Therefore, we may be concerned that the effect shown is not merely the result of electing a female councillor but rather of electing a councillor who is both less experienced and female.

Although we do not have data on candidate characteristics beyond name and gender, the panel structure of our data allows us to check for repeated appearances of the same name across elections. We construct variables that denote the number of times each candidate previously sought or won office in the past, and use these as proxies for experience. Therefore, our main specification shown in Table 3.3 includes these controls. Comparing Table 3.3 with Table A.4 shows little difference in the RD estimates before and after controlling for experience and city population.

Another potential explanation might be that the dip in per capita expenditure is explained not by council budget deliberations, but by exogenous changes in the city's revenue base. To address this concern, we estimate estimate the effect of a female win on total deficit, rather than total expenditure, and find a negative effect of around \$221 per capita (approximately 13% of the average total spending in our sample). The results are shown in Table 3.4. In other words, the effect persists even after accounting for revenue changes. As with the results on total spending, the level of statistical significance is somewhat sensitive to the specification used, but the estimates are consistently large and negative across all specifications. Binned scatter plots of revenue and residuals are shown in Figure 3.3.

What is the effect of a female win on the allocation of expenditure? Table 3.5 shows the results from running regressions for a number of expenditure categories that we were able to consistently identify across data sources. Note that these shares may not add to 1. In order to account for multiple testing, we test the joint hypothesis that the $\beta_1 = 0$ in all of the equations. (This is done under a Seemingly Unrelated Regressions (SUR) framework, which does not affect the point estimates as all our equations have the same right-hand side variables, but allows for cross-equation hypotheses to be tested). We find no evidence of any changes in expenditure composition, with or without control variables. We repeat the procedure with log per capita expenditure on each category instead of expenditure shares, and again find no evidence of a change in the way expenditure is allocated following a female win.

Next, we examine the effect of a close female win on the outcomes of the next election, particularly outcomes for female candidates. Table 3.6 shows no evidence that a female win impacts female participation or performance in the next election. We were slightly puzzled by the negative, albeit insignificant, estimates suggesting that a female win decreases female participation and success, but this effect goes away once we remove the current best loser from the count. Since the majority of seats turn over in two elections' time, when a woman wins she essentially removes herself from the next election, giving rise to the effect we observed. However, other women's participation and performance are not significantly affected.

It is interesting to consider the question of incumbency (dis)advantages in the context of city council elections. Table 3.7 shows the effect of a close female win on outcomes in the next two elections, in particular tracking the two candidates involved in the close gendered race. A female candidate who just won goes on to run within the next two elections 14pp to 17pp more than a female candidate who just lost, and wins 13pp to 15pp more. This indicates a significant incumbency effect for women. In contrast, male candidates in close gendered elections are not significantly affected by their incumbent status. This result recalls similar findings by previous papers and suggests a greater degree of voter learning concerning female office-holders. This is, however, not due to female candidates' inexperience, as the effect persists after controlling for candidates' experience. It is possible that voters have low ex-ante expectations relative to female candidates' true competence, although we cannot say for certain.

Still on Table 3.7, we find that a close female win increases the best loser's participation (but not their performance) in the next two elections. Note that, at the cutoff, the identity of the best loser switches from female to male. This means that men are less discouraged (or more encouraged) by a close defeat to participate next time, compared to women in a similar situation. Although this is pure conjecture, one possible explanation could be that women enter the race with a lower expectation of winning — as underdogs, so to speak. Once this expectation is confirmed, they are more likely to drop out. Men, on the other hand, are more likely to return next election. Table 3.8 shows similar results when we include all future elections.

3.6 Conclusion

This chapter investigates the effect of gender representation on policy at the municipal level in the United States. We show evidence that supports previous findings that gender representation does not affect the way municipal expenditure is allocated, or, if anything, decreases per capita expenditure. Our hypothesis is that changes in the gender composition of a council leads individuals to change their strategic behavior. Women, even if elected into the council, may find it difficult to express opinions that are at odds with the majority of the council; men, when faced with a larger minority, may push back against opposing preferences, thus leading to increased disagreement and a potential gridlock. We are excited about the next part of this project, which involves highly detailed data obtained from the records of proceedings of city council meetings. This will allow us to test our hypothesis directly and provide a more comprehensive understanding of gender in policymaking.

3.7 Figures and Tables

3.7.1 Figures

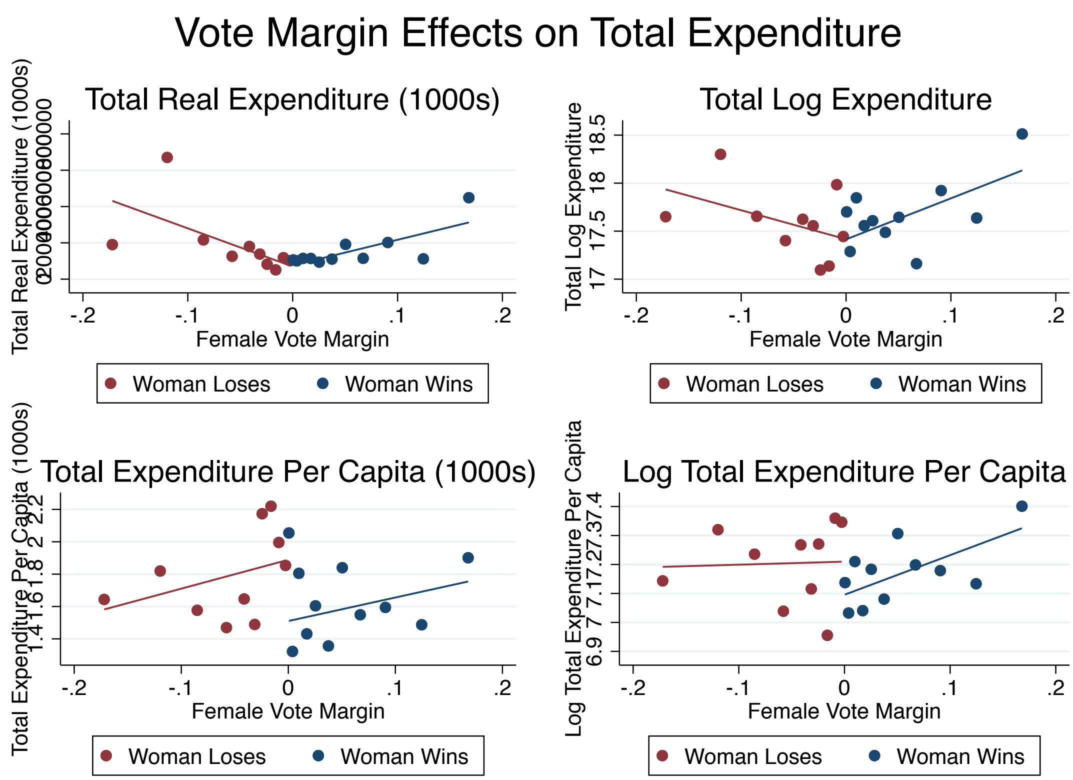


Figure 3·1: Total Expenditure Against Vote Margin

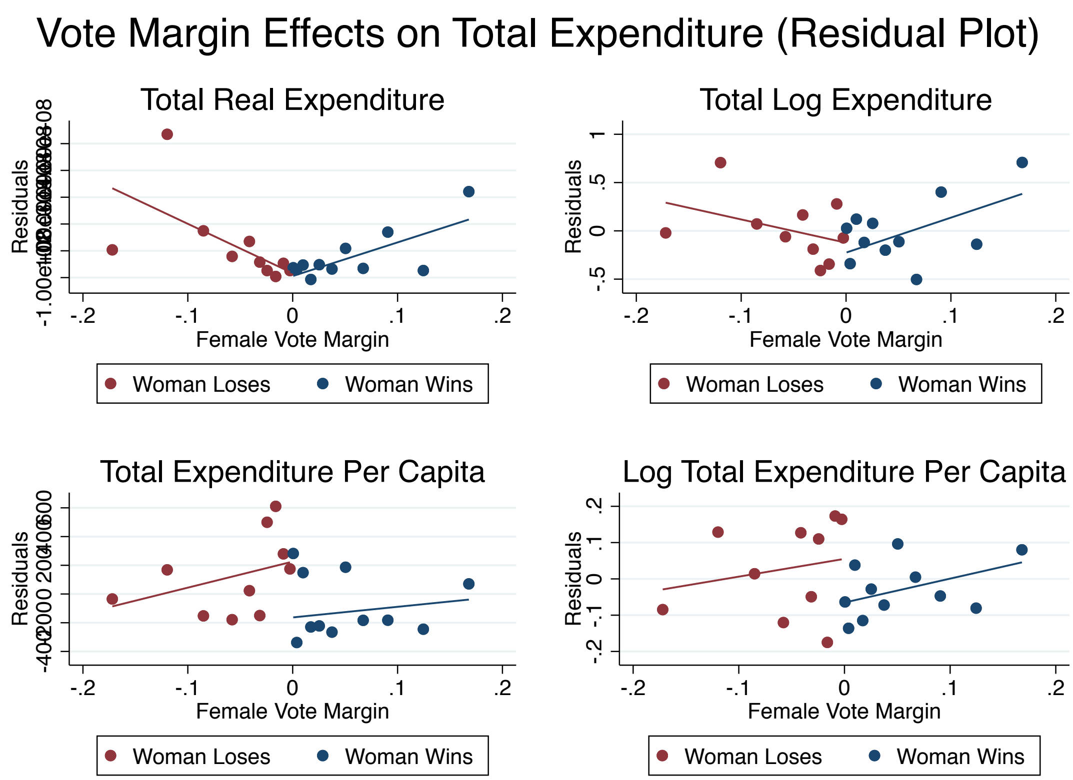


Figure 3·2: Total Expenditure Residuals (After Controlling for State and Year) Against Vote Margin

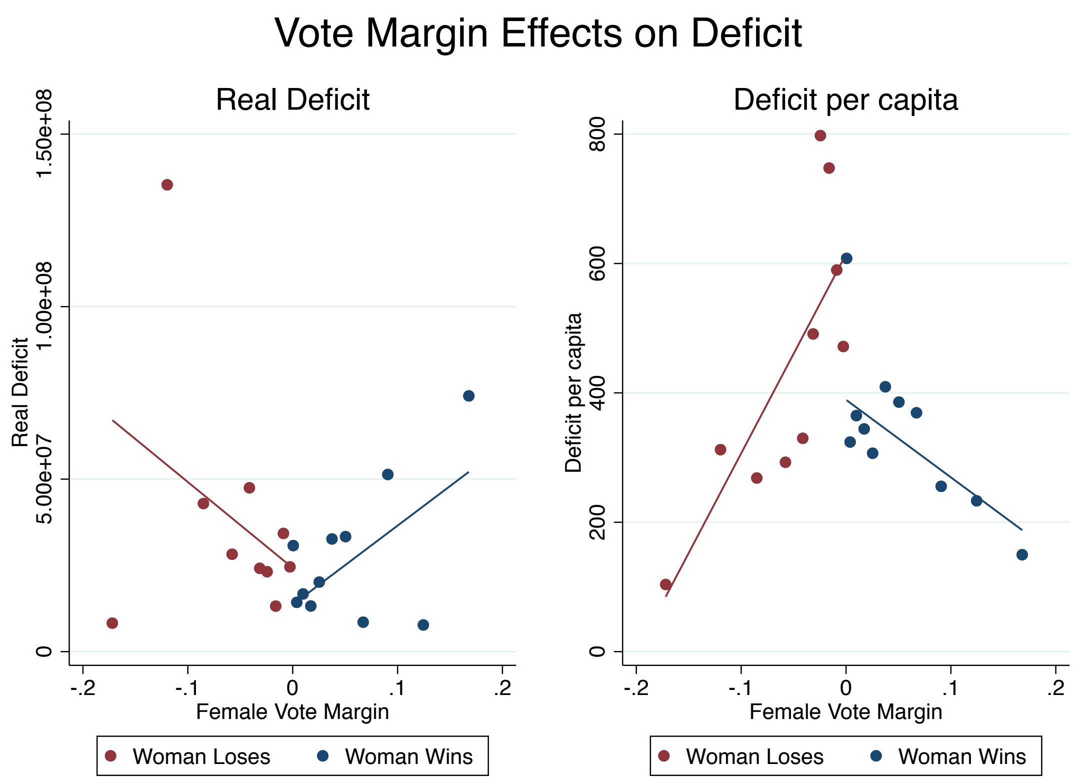


Figure 3·3: Deficit Against Vote Margin

3.7.2 Tables

Table 3.1: Election Data Summary Statistics

By Race	Number of Races	% Races Contested	% Races with Gender Contest	% Gender Contests Women Win	Mean Number of Seats	Mean Number of Candidates	Mean Votes per Seat
CEDA Data	2,746	0.904	0.287	0.496	2.00	4.52	9,297
From FOIA Requests	3,677	0.551	0.176	0.526	1.31	2.16	3,211
Combined	6,423	0.702	0.223	0.509	1.61	3.17	5,814
By Election	Number of Elections	% Any Contested Race	% Any Gender Contested Race		Mean Number of Races		
CEDA Data	2,233	0.941	0.326		1.23		
From FOIA Requests	1,277	0.791	0.381		2.88		
Combined	3,510	0.886	0.346		1.83		
By City	Number of Cities	% Ever Contested Race	% Ever Gender Contested Race		Mean Number of Races	Mean Number of Elections	Mean Years Covered
CEDA Data	498	0.996	0.773		5.51	4.48	7.52
From FOIA Requests	223	0.987	0.843		16.49	5.73	7.53
Combined	721	0.993	0.795		8.91	4.87	7.31

Table 3.2: 2012 Municipal Spending Summary Statistics

	All Cities			Cities with Election Data			Cities with Close Gendered Election		
	CA	Census	Combined	CA	Census	Combined	CA	Census	Combined
Total Spending (in millions)									
Total	126.83	25.15	27.61	110.13	89.09	103.50	131.72	111.82	125.58
Utilities	25.11	4.80	5.29	22.54	18.73	21.34	27.93	24.32	26.82
Parks and Recreation	5.27	0.98	1.09	5.07	4.93	5.02	5.71	6.48	5.95
Libraries	1.75	0.21	0.25	1.59	1.06	1.42	1.77	1.32	1.63
Air and Sea Ports	8.89	0.52	0.72	7.20	4.16	6.24	10.67	6.04	9.24
Police and Fire	28.01	3.97	4.55	26.83	17.12	23.77	31.35	21.19	28.21
Sewage and Waste	10.93	2.22	2.43	10.36	9.63	10.13	11.72	11.95	11.79
Roads and Parking	10.26	1.57	1.78	9.62	6.94	8.78	10.85	8.01	9.98
Health	5.17	1.21	1.31	0.81	1.56	1.05	1.11	1.97	1.37
Spending Per Capita									
Total	4,569.67	1,647.24	1,717.70	4,578.36	1,926.24	3,742.45	1,582.93	1,841.89	1,662.91
Utilities	2,198.92	412.22	455.30	2,213.72	403.15	1,643.05	218.06	359.42	261.72
Parks and Recreation	100.85	60.19	61.17	100.94	101.68	101.17	98.49	116.17	103.95
Libraries	19.14	11.07	11.26	19.10	29.74	22.46	21.58	28.52	23.72
Air and Sea Ports	17.12	11.03	11.18	15.22	15.80	15.41	20.94	17.72	19.95
Police and Fire	691.57	238.59	249.51	693.70	311.86	573.35	413.13	303.55	379.28
Sewage and Waste	238.08	220.10	220.53	238.27	205.01	227.79	197.34	189.82	195.02
Roads and Parking	218.55	145.26	147.02	216.80	206.39	213.52	171.42	209.82	183.28
Health	39.50	39.62	39.61	34.75	86.14	50.94	14.51	74.44	33.02
Number of Cities	482	19,513	19,995	478	220	698	320	143	463

Table 3.3: RD Estimates: Effect of Close Female Win on Expenditure, Controlling for Population and Winner's Experience

	Total Real Expenditure			Total Log Expenditure		
	Optimal Bandwidth	Half Optimal Bandwidth	Twice Optimal Bandwidth	Optimal Bandwidth	Half Optimal Bandwidth	Twice Optimal Bandwidth
Female Winner	-2105423.9	-8748458.1	15731494.6	0.120	0.0764	-0.0694
	(14805855.5)	(20866692.8)	(13368402.3)	(0.134)	(0.160)	(0.119)
Bandwidth	0.0660	0.0330	0.132	0.127	0.0635	0.254
	Total Expenditure Per Capita			**Log Total Expenditure Per Capita**		
	Optimal Bandwidth	Half Optimal Bandwidth	Twice Optimal Bandwidth	Optimal Bandwidth	Half Optimal Bandwidth	Twice Optimal Bandwidth
Female Winner	-393.2	-624.8*	-682.3*	-0.101	-0.192**	-0.111**
	(268.8)	(334.1)	(370.6)	(0.0655)	(0.0826)	(0.0536)
Bandwidth	0.119	0.0595	0.238	0.164	0.0820	0.328

Standard errors are clustered at the city level and are in parentheses. All regressions include year and state fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.4: RD Estimates: Effect of Close Female Win on Deficit, Controlling for Population and Winner's Experience

	Total Real Deficit (000s)			Deficit Per Capita		
	Optimal Bandwidth	Half Optimal Bandwidth	Twice Optimal Bandwidth	Optimal Bandwidth	Half Optimal Bandwidth	Twice Optimal Bandwidth
Female Winner	-7,090.5*	-12,262.2**	-1,421.0	-221.3*	-329.6**	-243.1*
	(4266.1)	(5662.5)	(3720.5)	(126.6)	(143.9)	(127.3)
Bandwidth	0.0660	0.0330	0.132	0.119	0.0595	0.238

Standard errors are clustered at the city level and are in parentheses. All regressions include year and state fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.5: RD Estimates: Effect of Close Female Win on Expenditure Composition, Controlling for Population and Winner's Experience

Bandwidth	0.100	0.150	0.200	0.250
Dependent Variable:				
% Public Utility Expenditure	-0.0145	-0.0152	-0.0131	-0.0141
	(0.0124)	(0.0114)	(0.0102)	(0.00928)
% Health and Hospital Expenditure	0.00344	0.00695	0.00479	0.00474
	(0.0142)	(0.0121)	(0.0113)	(0.00993)
% Parks and Recreation Expenditure	0.000360	0.00356	0.00385	0.00252
	(0.00629)	(0.00544)	(0.00490)	(0.00452)
% Library Expenditure	-0.00103	-0.000925	-0.00216	-0.00210
	(0.00232)	(0.00194)	(0.00169)	(0.00152)
% Housing and Community Dev Expenditure	-0.00241	-0.00528	-0.00552	-0.00255
	(0.00785)	(0.00678)	(0.00617)	(0.00594)
% Airports and Water Ports Expenditure	-0.000253	-0.00111	-0.00605**	-0.00647*
	(0.00258)	(0.00313)	(0.00301)	(0.00358)
% Police and Fire Expenditure	0.0125	0.00675	0.00715	0.0105
	(0.0127)	(0.0112)	(0.0100)	(0.00906)
% Sewerage and Waste Expenditure	0.00677	0.00348	0.00366	0.00476
	(0.0129)	(0.0118)	(0.0105)	(0.00936)
% Correctional Expenditure	0.000189	-0.000754	-0.000658	-0.000360
	(0.00176)	(0.000971)	(0.000872)	(0.000815)
% Roads and Parking Expenditure	0.0119	0.0115	0.0104	0.00802
	(0.00886)	(0.00831)	(0.00758)	(0.00722)
Joint test	8.304	7.321	10.96	10.96
p-value	0.599	0.695	0.361	0.361

Standard errors are clustered at the city level and are in parentheses.
All regressions include year and state fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.6: RD Estimates: Effect of Close Female Win on Next Election

Bandwidth	0.100	0.150	0.200	0.250
Dependent Variable:				
Number of Candidates Next Election	-0.0477	-0.224	-0.255	-0.264
	(0.423)	(0.400)	(0.332)	(0.303)
Number of Female Candidates Next Election	-0.199	-0.256	-0.217	-0.242
	(0.219)	(0.207)	(0.179)	(0.165)
Number of Winners Next Election	-0.00169	-0.0270	-0.0171	-0.0244
	(0.138)	(0.137)	(0.113)	(0.103)
Number of Female Wins Next Election	-0.141	-0.144	-0.127	-0.140
	(0.113)	(0.123)	(0.100)	(0.0879)
Number of Female Candidates Next Election	0.00969	-0.0749	-0.0681	-0.0909
Besides Current Best Loser	(0.217)	(0.205)	(0.179)	(0.165)
Number of Female Wins Next Election	-0.0368	-0.0536	-0.0429	-0.0569
Besides Current Best Loser	(0.117)	(0.123)	(0.102)	(0.0900)

Standard errors are clustered at the city level and are in parentheses.
All regressions include year and state fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7: RD Estimates: Effect of Close Female Win on Next Two Elections, Controlling for Candidates' Experience

Bandwidth	0.100	0.150	0.200	0.250
Dependent Variable:				
Best loser runs at least once in next 2 elections	0.0802*	0.0566	0.0802**	0.0688*
	(0.0480)	(0.0413)	(0.0379)	(0.0352)
Worst winner runs at least once in next 2 elections	0.0726	0.109***	0.107***	0.0968***
	(0.0461)	(0.0409)	(0.0377)	(0.0351)
Male candidate runs at least once in next 2 elections	0.0131	0.0145	0.0132	-0.00842
	(0.0477)	(0.0423)	(0.0393)	(0.0373)
Female candidate runs at least once in next 2 elections	0.140***	0.151***	0.174***	0.174***
	(0.0461)	(0.0404)	(0.0381)	(0.0351)
Best loser wins at least once in next 2 elections	0.0378	0.0243	0.0246	0.0202
	(0.0407)	(0.0348)	(0.0321)	(0.0300)
Worst winner wins at least once in next 2 elections	0.0850*	0.117***	0.101***	0.0946***
	(0.0450)	(0.0400)	(0.0367)	(0.0345)
Male candidate wins at least once in next 2 elections	-0.00782	0.000942	-0.0139	-0.0399
	(0.0431)	(0.0377)	(0.0350)	(0.0327)
Female candidate wins at least once in next 2 elections	0.131***	0.141***	0.140***	0.155***
	(0.0418)	(0.0367)	(0.0337)	(0.0312)

Standard errors are clustered at the city level and are in parentheses.
All regressions include year and state fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.8: RD Estimates: Effect of Close Female Win on All Future Elections, Controlling for Candidates' Experience

Bandwidth	0.100	0.150	0.200	0.250
Dependent Variable:				
Number of times the best loser runs in the future	0.110	0.0858	0.118**	0.113**
	(0.0743)	(0.0637)	(0.0575)	(0.0527)
Number of future wins by the best loser	0.0284	0.0115	0.0248	0.0239
	(0.0565)	(0.0493)	(0.0455)	(0.0421)
Number of times the worst winner runs in the future	0.111*	0.113**	0.100**	0.117**
	(0.0589)	(0.0531)	(0.0507)	(0.0491)
Number of future wins by the worst winner	0.111*	0.123**	0.0979**	0.0993**
	(0.0586)	(0.0525)	(0.0490)	(0.0472)
Number of future runs by the male candidate	0.0841	0.0609	0.0622	0.0543
	(0.0685)	(0.0603)	(0.0568)	(0.0531)
Number of future wins by the male candidate	0.0185	0.0202	0.00204	-0.0192
	(0.0596)	(0.0526)	(0.0490)	(0.0449)
Number of future runs by the female candidate	0.136**	0.138**	0.156***	0.176***
	(0.0648)	(0.0574)	(0.0534)	(0.0493)
Number of future wins by the female candidate	0.121**	0.115**	0.121**	0.142***
	(0.0553)	(0.0506)	(0.0475)	(0.0441)
Future wins conditional on running by the best loser	-0.00556	-0.0669	-0.106	-0.106
	(0.113)	(0.101)	(0.0876)	(0.0852)
Future wins conditional on running by the worst winner	-0.00836	0.0507	0.0221	0.0274
	(0.0924)	(0.0760)	(0.0683)	(0.0643)
Future wins conditional on running by the male candidate	-0.0937	-0.131	-0.194**	-0.224***
	(0.105)	(0.0867)	(0.0839)	(0.0759)
Future wins conditional on running by the female candidate	0.126	0.119	0.104	0.148**
	(0.0975)	(0.0843)	(0.0764)	(0.0728)

Standard errors are clustered at the city level and are in parentheses. All regressions include year and state fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Appendix A

Appendices

A.1 Chapter 1: More results

A.1.1 Moderately Biased Sender

The model in Section 1 illustrates that political competition can generate endogenous credibility for an extremely biased sender. In this section, I allow the sender's bias to be moderate, as in Crawford and Sobel (1982). I show that in some cases, the equilibria in Section 1 are also equilibria for smaller bias. In particular, when the prior is uniform, the set of equilibria under extreme bias is a subset of equilibria under moderate bias.

Lemma 6. Consider the game in Section 1 with the following modification:

$$U_S(a^1, a^2; \theta) = -E((a^w - \theta - b)^2 \,\Big|\, a^1, a^2; \theta)$$

Also, let $\boldsymbol{\sigma}$ be an equilibrium of the original game, such that the set of induceable policies is has at most two elements, i.e. $|A_R| \leq 2$, where $A_R = \{a_R : \exists\,\theta\,\text{s.t. } a_R = \sigma_R(\sigma_{SR}(\theta))\}$. Then $\boldsymbol{\sigma}$ is also an equilibrium of the modified game with $b > 0$.

Corollary. If $\theta \sim Uniform[0, 1]$, then any equilibrium of the original game is also an equilibrium of the modified game with moderate bias.

A.1.2 Imperfectly Informed Sender

So far, I have assumed that the sender has perfect knowledge of the median voter's preference. However, in reality it is doubtful that anyone could perfectly predict the outcome of an election. In this subsection, I assume that the sender observes an imperfect signal about θ.

Suppose there is a small probability that the sender's signal is entirely uninformative. For instance, suppose that the sender observes the true state with probability $1 - \varepsilon$, but with probability ε the sender receives a meaningless signal that is uniformly drawn from the state space. Then the only equilibrium is babbling.

To see why, remember that the sender's credibility arises from pragmatism: there is no point in making a false recommendation if that policy will be defeated for sure. However, if the sender's information is imprecise, any policy could win with a small probability! Even if the sender's information suggests that the status quo will win, there is always a small chance that an exaggerated recommendation will bear fruit. This creates a small misalignment between the sender and receiver: the sender would like to recommend a high policy as long as it has any chance at all of winning, whereas the candidate is more cautious and prefers policies that are most likely to win.

This misalignment is so small, however, that small modifications to the model will eliminate it. Below, I consider one such modification: a small cost of sending a message. This assumption is entirely plausible in real world situations. In my model, even a small cost is enough to create credibility, even though the sender is extremely biased (has state-independent preferences). This is because most of the sender's credibility already arises endogenously.

Senders with more precise information have more credibility in equilibrium. If the sender is perfectly informed, nothing else is needed. However, the more imprecise the sender's information, the more important other considerations will be for credibility.

This is in contrast with Crawford and Sobel (1982), where the precision of the sender's information does not matter for credibility (because agents' best responses only depend on the posterior mean after each signal). My model predicts that special interests with precise information need not do much to convince politicians, whereas equally biased senders with imprecise signals need to go to greater lengths to convey similar information. Likewise, interests with imprecise information must be punished *harder* when their recommendations lead to a loss.

Messages are *almost* costless Suppose that, instead of being completely costless, all messages have a small cost $c > 0$ to the sender. However, there is one message, $m_0 \in M$, that costs nothing — interpret that as "no message". Additionally, assume that the candidate has a small bias towards high policies.

First, if the sender is perfectly informed, the equilibria that remain are exactly the interval equilibria. Furthermore, "no message" must signify a recommendation for the status quo (lowest) policy.

If instead the sender receives an almost perfect signal, i.e. $\varepsilon \to 0$, the incentive to exaggerate is outweighed by the cost of the message. The sender knows that the high policy could win with non-zero probability, but that event is so unlikely that it is not worth sending a message. For the candidate, receiving the message means that a high policy is very likely to win, however there is still a small chance that the sender is wrong, and the status quo might be what voters truly prefer. If the candidate only cared about winning, they would deviate and choose a policy closer to the status quo policy (instead of the high policy), just in case the sender was wrong. However, a small upward bias will make sure they follow the recommendation.

Lemma 5. Suppose that $\boldsymbol{\sigma}$ is an interval equilibrium of the original cheap-talk game. Then with each of the following modifications, $\boldsymbol{\sigma}$ is still an equilibrium, as long as $\sigma_R(m_0) = a_U$ and $c > 0$ is small.

1. Small cost of sending a message, and small candidate bias.

$$U_S(\boldsymbol{\sigma};\theta) = Pr(1 \text{ wins})\ \ u_S(a^1)\ \ +\ \ Pr(2 \text{ wins})\ \ u_S(a^2) - c\ \sum_i \nVdash\{m^i \neq m_0\}$$

$$= Pr(w = 1 \,|\, \boldsymbol{\sigma};\theta)\ \ u_S(\sigma_1(\sigma_{S1}(\theta)))\ \ +\ \ Pr(w = 2 \,|\, \boldsymbol{\sigma};\theta)\ \ u_S(\sigma_2(\sigma_{S2}(\theta)))$$

$$- c\ \sum_i \mathbf{1}\{\sigma_{Si}(\theta) \neq m_0\}$$

$$U_R(\boldsymbol{\sigma};\theta) = Pr(w = R \,|\, \boldsymbol{\sigma};\theta) + \beta\ \sigma_i(\sigma_{SR}(\theta)), \quad \text{where } \beta > 0 \text{ is small.}$$

2. Small cost of sending a message, small candidate bias, and imperfectly informed sender.

   Preferences as stated above, and:

   S observes a signal $s = \begin{cases} \theta & \text{w.p. } 1 - \varepsilon, \\ x \sim Uniform[0,1] & \text{w.p. } \varepsilon. \end{cases}$

   (S's strategy is now a function of s instead of θ).

A.1.3 Two senders with Opposite Biases

Suppose that, instead of one extremely biased sender, there were two senders with diametrically opposed preferences (increasing and decreasing) in policy. Then, it is trivially true that perfect revelation by both senders is supported by an equilibrium. Consider the following strategy profile: S_1 and S_2 each perfectly reveal the state. Candidate 1 always chooses a policy equal to the state revealed by S_1 and Candidate 2 likewise listens to S_2. Then neither sender has an incentive to deviate, as there is always at least one candidate who chooses a policy equal to the median voter's bliss point. Therefore that policy is guaranteed to win. Neither candidate has an incentive to deviate, as deviating guarantees a loss.

This is akin to the ideal confrontational system which is at the basis of most democratic institutions, coming back to the Downsian idea that self-interested political actors end up serving the interests of the median voter. The result, however, rests on a number of strict assumptions, chiefly that both senders have access to the same precise information. The model presented in Chapter One, on the other hand, is more applicable whenever the special interests on one side of the argument have a disproportionate informational advantage over the other. In the real world, this arises naturally due to differences in wealth, expertise, or the ability to mobilise resources.

A.2 Proofs

A.2.1 Chapter 1

Proposition 1: No public messages. If we restrict the sender's strategy to public messages (i.e. $\sigma_{S1} = \sigma_{S2}$), then the only equilibrium is babbling, and both candidates always choose θ_m, the median of F, regardless of messages.

Section 2.2 already has an informal proof of Proposition 1. That proof is valid if the median given each posterior is unique. This need not be the case, however, if the posterior is not continuous - in particular, if it assigns zero probability to middling values. Here is a formal proof that takes this into account.

The intuition is as follows: for small θ, only the lower policy (between the two candidates) matters, so S always induces the highest such policy. Therefore, the lower policy is unique. For some moderate θ, only the higher policy (between the two candidates) matters, and so S always induces the highest such policy, which means the higher and lower policy cannot both be posterior medians unless they are equal.

Fix an equilibrium. By the median voter theorem, each candidate must, given each message, choose a policy platform among the set of medians given the posterior, and win with probability half. Formally,

$$\text{For each } i \in \{1, 2\} \text{ and for each } m \in M, \quad \sigma_i(m) \in \theta_m(m)$$

where $\theta_m(m) \equiv \left\{\hat{\theta} \in \Theta : \, Pr(\theta \leq \hat{\theta} \,|\, m) \geq \frac{1}{2} \text{ and } Pr(\theta \geq \hat{\theta} \,|\, m) \geq \frac{1}{2}\right\}$.

Now define the set of lower policies

$$A_\ell = \{a \in [0, 1] : a = \min_i \sigma_i(m) \text{ for some } m\}$$

Now let $\bar{a}_\ell = \sup A_\ell$. Fix any $a \in A_\ell \backslash \{\bar{a}_\ell\}$. For all $\theta \leq a$, the sender strictly prefers to induce (something arbitrarily close to) $\bar{a}_\ell$ rather than a, because in both cases the lower policy (between the two candidates) wins for sure. But then $Pr(\theta \leq a|\, a \text{ is induced}) = 0$ and therefore a cannot be the median of the posterior given the message that induces a.

Therefore, the only possibility is that $A_\ell = \{a_\ell\}$ is a singleton. This means that, after any message, at least one of the candidates must choose a_ℓ. We need to show that both candidates always choose a_ℓ.

Suppose not. Define $a_h = \sup\{a \in [0,1] : a = \sigma_i(m) \text{ for some } m, i\}$ to be the highest policy on the equilibrium path. This policy must be a maximum, otherwise when $\theta = 1$, the sender's optimal message does not exist. For all $\theta \in (\frac{a_\ell + a_h}{2}, a_h)$, the sender must induce a_h since it will win for sure. But this means that a_ℓ and a_h cannot both be medians of the posterior given the message that induces them. Contradiction.

The last step is to show that the unique policy chosen, a, must be equal to $\theta_{\frac{1}{2}}$, the median of F. Suppose not, then given some message the posterior must place unequal mass above and below a. Then each candidate can deviate by choosing instead a median of the posterior, winning with probability strictly higher than half.

Proposition 2: Babble to one, talk to the other. On the equilibrium path, at least one candidate's strategy must be independent of messages, i.e. $\exists U \in \{1,2\}$ such that $\sigma_U(\sigma_{SU}(\theta)) = a_U \quad \forall \theta$.

Suppose, by contradiction, that the sender induces at least two distinct policy platforms, $a_i^\ell < a_i^h$, for each of the candidates in some equilibrium. W.l.o.g., let $a_i^h = \sup A_i$, where $A_i = \{a_i : \exists\, \theta \,\text{s.t.}\; a_i = \sigma_i(\sigma_{Si}(\theta))\}$.

First, a_i^h must be induced for some θ and some i (i.e. a_i^h is a max, not just a sup, for some i), otherwise the sender's best response does not exist when $\theta \geq \max\{a_1^h, a_2^h\}$. Fix that i. For $j \neq i$, the sender induces a_j^ℓ only if it loses to a_i^h. (If for some θ, a_j^ℓ has non-zero chance of winning, it is strictly better for the sender to induce a_i^h

and something arbitrarily close to a_j^h.) Then after receiving message a_j^ℓ, candidate j knows that i is playing a_i^h and that a_j^ℓ must lose for sure. So j can deviate profitably to playing a_i^h and winning with probability $\frac{1}{2}$.

Proposition 3: Uninformed candidate competes at lower end. Let U denote the candidate whose strategy is independent of messages, and let R denote the other. Then $a_U = \inf A_R, \quad \text{where } A_R = \{a_R : \exists\, \theta \text{ s.t. } a_R = \sigma_R(\sigma_{SR}(\theta))\}$.

Denote $a_R^\ell \equiv \inf \{a_R : \exists\, \theta \text{ s.t. } a_R = \sigma_R(\sigma_{SR}(\theta))\}$.

Suppose $a_U < a_R^\ell$. Then for any $\theta > \frac{a_U + a_R^\ell}{2}$, R must win, otherwise the sender is not playing best response. Then U can deviate profitably by choosing $a'_U = a_R^\ell - \varepsilon$, where ε is arbitrarily small.

Suppose $a_U > a_R^\ell$. Then there exist some θ for which the sender induces $a_R < a_U$. Fix such an a_R. If given a_R is induced, R never wins, then R can do strictly better by choosing a_U. If R sometimes wins, then the sender can deviate by inducing something strictly higher than a_R when a_R would have won. (There must exist at least two inducible actions for R, if $a_U > a_R^\ell$, because otherwise the candidates are not playing best response).

Proposition 4. Uninformed policy no higher than the median of F. $a_U \leq \theta_{\frac{1}{2}}$.

Suppose $a_U > \theta_{\frac{1}{2}}$. First note that a_U must be a median of the posterior conditional on the sender recommending a_U to R, otherwise R could deviate by a small amount and win with probability strictly greater than 0.5. Now there must be some $a_R \in A_R \backslash \{a_U\}$ such that the posterior (given S recommends a_R) assigns greater than half probability to $\theta < a_U$. Since $a_R > a_U$ by Proposition 3, R wins less than half of the time, and R can deviate to some policy below a_U and win strictly more than half the time.

Proposition 5: Finitely many messages. Assume F is a non-doctrinaire prior, i.e. $f(\theta) > 0$ for all $\theta \in (0,1)$. The set of policy platforms for R that are induced in equilibrium, $A_R = \{a_R : \exists\, \theta \,\text{s.t.}\; a_R = \sigma_R(\sigma_{SR}(\theta))\}$, is finite.

Note: $f(\theta) > 0$ for $\theta \in (0,1)$ is stronger than full support, but is satisfied under full support and single-peakedness. Note that $f(\theta)$ is allowed to be 0 at the extremes.

First we make the following claim.

Claim: For all $a > a_U$,

$$\sup\{a_R \in A_R : a_R < a\} < a \quad \text{and is a } \max .$$

i.e. the highest element that is strictly smaller than a exists.

First, suppose $\sup\{a_R \in A_R : a_R < a\} = a$ for some a. Then when $\theta = \frac{a_U + a}{2}$, S's optimal strategy does not exist, because S can deviate to inducing a policy that is arbitrarily close below a that can win for sure. Now, suppose that the sup is not a max. Then set $a' = \sup\{a_R \in A_R : a_R < a\}$ and apply the first part of the claim. Contradiction.

Suppose A_R is infinite. Construct a series of arbitrarily small upward deviations for U, the uninformed candidate, as follows.

Define the function $h : (a_U, 1] \to A_R \backslash \{a_U\}$ by saying $h(a) = \max\{a_R \in A_R : a_R < a\}$. Starting with $a_0 = h(1)$, let $a_t = h(a_{t-1})$ for all t. Since A_R is infinite, this gives us an infinite series of strictly decreasing elements, from the highest element of A_R to the lowest. Now let $\epsilon_t = \frac{1}{2} \min_{1 \leq k \leq t}\{a_{t-1} - a_t\} > 0$.

By deviating upwards to $a_U + \epsilon_t$, U loses approximately $\epsilon_t \frac{1}{2} f(a_U)$, which occurs when θ happens to be just below $\frac{a_U + a_U + \epsilon}{2}$. However, U gains at least $\epsilon_t \sum_{k=1}^{t} f(\frac{a_k + a_U}{2})$. (Remember, at θ just above $\frac{a_k + a_U}{2}$, R would usually just win, but now U wins instead). As $\epsilon_t \to 0$, this sum grows to infinity, and the deviation is profitable.

Lemma 1 If $\theta \sim Uniform[0,1]$, then A_R is at most a doubleton. This means that in equilibrium, the sender recommends one out of a set of at most two policy platforms to the informed candidate.

Suppose not, i.e. $|A_R| \geq 3$. Then U can deviate by choosing $a'_U = a_U + \varepsilon$ instead. From Proposition 3 and 5, we know that $a_U = \min A_R$. Also, a_U is the median of the posterior on θ conditional on R choosing a_U (i.e. conditional on $\sigma_R(\sigma_{SR}(\theta)) = a_U$), by a reasoning similar to the median voter theorem. As $\epsilon \to 0$, U's gain from deviating converges to:

$$\varepsilon\Big(-\frac{1}{2}f(a_U) + \frac{1}{2}\sum_{a_R \in A_R \backslash \{a_U\}} f(\frac{a_U + a_R}{2})\Big) \quad = \quad \varepsilon\Big(-\frac{1}{2} + \frac{1}{2}(|A_R| - 1)\Big) > 0$$

The first term in brackets refers to the loss from deviating when R chooses a_U and $\theta < \frac{a_U + a'_U}{2}$. The second term refers to the gain when $\sigma_R(\sigma_{SR}(\theta)) \neq a_U$. The sender recommends $a_R \in A_R \backslash \{a_U\}$ whenever $\theta > \frac{a_U + a_R}{2}$, i.e. whenever a_R wins against a_U. So for U, deviating by ε means pushing the cutoff for a win by $\frac{1}{2}\varepsilon$ and capturing those θ between the old and new cutoffs. There is a cutoff for each $a_R \in A_R \backslash \{a_U\}$, so add up the gains from moving all of them.

Lemma 2. Interval equilibria. If $\theta \sim Uniform[0,1]$, the following is an equilibrium iff $a_\ell \in [\frac{1}{4}, \frac{1}{3}]$ and $a_h = 3a_\ell$.

Rename the candidates $R, U \in \{1, 2\}$, where $R \neq U$, to denote who receives information (R) and who receives a babbling message (U). Fix some arbitrary $m_U \in M$. Also fix a partition $\{M_\ell, M_h\}$ of M, and arbitrary elements $m_\ell \in M_\ell$ and $m_h \in M_h$.

$$\sigma_S(\theta) = \begin{cases} (m_U, m_\ell) & \text{if } \theta < \frac{a_\ell + a_h}{2}, \\ (m_U, m_h) & \text{otherwise.} \end{cases}$$

$$\sigma_U(m) = a_U = a_\ell \quad \text{for all } m \in M, \quad \text{with prior belief after every } m.$$

$$\sigma_R(m) = \begin{cases} a_\ell & \text{for all } m \in M_\ell, \quad \text{with belief } \theta \sim Uniform[0, \frac{a_\ell + a_h}{2}], \\ a_h & \text{for all } m \in M_h, \quad \text{with belief } \theta \sim Uniform[\frac{a_\ell + a_h}{2}, 1]. \end{cases}$$

It is easy to show that each player is playing best response given others' strategies. When $\theta < \frac{a_\ell + a_h}{2}$, S must choose between inducing $a_R = a_\ell$ or $a_R = a_h$. Since a_h cannot win anyway, S is indifferent between all messages. When $\theta \geq \frac{a_\ell + a_h}{2}$, a_h can win so it is strictly optimal to induce a_h.

a_ℓ is optimal for U because any other policy gives a strictly (weakly if $a_\ell = \frac{1}{4}$) lower probability of winning. For R, there are two possible posterior beliefs about θ. After receiving message m_ℓ, the posterior is a uniform distribution over $[0, \frac{a_\ell + a_h}{2})$, and a_U is the median of that posterior, so just like under Hotelling-Downs, playing the median a_U is optimal. After receiving message m_h, R knows $\theta > \frac{a_\ell + a_h}{2}$, so a_h guarantees a win with probability 1. So it is optimal.

Lemma 3. Sender-optimal equilibrium. If $\theta \sim Uniform[0, 1]$, then the following is an equilibrium, and it achieves the highest equilibrium expected payoff for the sender.

Fix some arbitrary $m_U \in M$. Also fix a partition $\{M_1, M_2\}$ of M, and arbitrary elements $m_1 \in M_1$ and $m_2 \in M_2$.

$$\sigma_S(\theta) = \begin{cases} (m_U, m_1) & \text{if } \theta \in [0.25, 0.75), \\ (m_U, m_2) & \text{otherwise.} \end{cases}$$

$$\sigma_U(m) = 0.5 \quad \text{for all } m \in M, \quad \text{with prior belief after every } m.$$

$$\sigma_R(m) = \begin{cases} 0.5 & \text{for all } m \in M_1, \quad \text{with belief } \theta \sim Uniform[0.25, 0.75), \\ 1 & \text{for all } m \in M_2, \quad \text{with posterior given } \theta \in [0, 0.25) \cup [0.75, 1]. \end{cases}$$

If u_S is linear, this equilibrium gives the sender expected utility $E_\theta(U_S(\boldsymbol{\sigma}^*; \theta)) = 0.625$.

It is easy to show that the above is an equilibrium. To show that it is the best equilibrium for the sender, we refer to Lemma 1 which implies that that any informative equilibrium has exactly two induceable policies for R. Since, in equilibrium, each induceable policy wins iff the median voter prefers it to the other, we can easily show that the sender's utility is increasing in both induceable policies. In the above equilibrium, the higher induceable policy is as high as possible at 1. Then it suffices to show that, no matter what the higher induceable policy is, the lower policy $\min A_R = a_U \leq \frac{1}{2}$ in any equilibrium.

Suppose that $a_U > \frac{1}{2}$ in some equilibrium. First, a_U must be the median of R's posterior after receiving a recommendation to choose policy a_U. (Otherwise, R would deviate). Also, name the other induceable policy a_R, where $a_R > a_U$ by Proposition 3. Then for R to follow a recommendation of a_R, it must be that $Pr(\theta > a_U | \sigma_{SR}(\theta) = a_R) \geq \frac{1}{2}$, otherwise it's better for R to deviate to $a_U - \varepsilon$. Combining these two statements, we must have that $Pr(\theta > a_U) \geq \frac{1}{2}$, which under Uniform distribution means $a_U \leq \frac{1}{2}$.

Finally, note that any equilibrium where the set of induceable policies is $\{\frac{1}{2}, 1\}$ is best for the sender.

Lemma 4. Welfare If $\theta \sim Uniform[0, 1]$ and u_S is linear, then any non-babbling equilibrium gives weakly higher utility to the sender and the informed receiver, and strictly higher utility to the median voter, than the babbling equilibrium.

By Lemma 1, a non-babbling equilibrium has exactly two on-the-equilibrium-path policies, call them a_1, $a_2 \in [0,1]$. W.l.o.g. let $a_1 > a_2$. By Proposition 4, $a_2 < \frac{1}{2}$. Furthermore, $a_1 > \frac{1}{2}$ because otherwise U would deviate to $\frac{1}{2}$. Also note that a_1 is implemented if and only if $\theta \geq \frac{a_1+a_2}{2}$.

Under babbling, the implemented policy is constant at $\frac{1}{2}$. This gives utilities $\frac{1}{2}$, $\frac{1}{2}$ and $-\frac{1}{12}$ to the sender, informed receiver and median voter respectively.

Under a non-babbling equilibrium, the sender's utility is no lower than $\frac{1}{2}$ because $\frac{a_1+a_2}{2} \geq \frac{1}{2}$. (Suppose $\frac{a_1+a_2}{2} < \frac{1}{2}$, then U could deviate profitably to slightly above a_2, which increases U's winning probability from a_2 to $1-a_1$). The informed receiver's utility is at least $\frac{1}{2}$ because otherwise, he can deviate to a constant policy of $\frac{1}{2}$, which guarantees him a utility of at least $\frac{1}{2}$. That the median voter's utility is strictly higher than $\frac{1}{2}$ can be shown by simple algebra.

Proposition 6a. Babble if f is strictly increasing. If the prior density f is strictly increasing in θ, babbling is the only equilibrium.

Fix any non-babbling equilibrium. For all $a_R \in A_R \backslash \{a_U\}$,

$$f(\frac{a_U + a_R}{2}) > f(a_U)$$

Now consider a small deviation for U, where U chooses $a_U + \varepsilon$ instead of a_U. As $\varepsilon \to 0$, the gain from deviating converges to

$$\varepsilon\Big(-\frac{1}{2}f(a_U) + \frac{1}{2}\sum_{a_R \in A_R \backslash \{a_U\}} f(\frac{a_U + a_R}{2})\Big) > 0$$

Proposition 6b. Non-babbling equilibrium if f is weakly decreasing. If the prior density f is decreasing in θ, a non-babbling equilibrium exists.

The proof is constructive. It is easy to show that there exists an interval equilibrium with two policies, $a_\ell = median(\theta \mid \theta < \frac{1}{2})$ and $a_h = 1 - a_\ell$, where S reveals whether θ is above or below $\frac{1}{2}$.

Proposition 7. Symmetric single-peaked prior. Suppose the prior f is single-peaked and symmetric around $\theta = \frac{1}{2}$. Then a non-babbling equilibrium exists.

Again, the proof is constructive. It is easy to show that there exists an equilibrium with two policies and two cutoffs. Let $x = median(\theta \mid \theta < \frac{1}{2}) \geq \frac{1}{4}$, and let the policies be $a_e ll = \frac{1}{2}$ and $a_h = 2(1-x) - \frac{1}{2}$.

S reveals whether θ is in the interval $[x, 1-x)$. If it is, R chooses $a_R = \frac{1}{2}$, otherwise $a_R = a_h$.

A.2.2 Chapter 2

Proposition 8. Sender's optimum. The following signal structure is the solution to the sender's public persuasion problem, and achieves the maximal utility equal to $E(u_S(\theta)\,|\,\theta > \theta_{\frac{1}{2}})$.

$$p(m|\theta) = 1 \quad \text{iff} \quad \theta = F^{-1}(\frac{1}{2} \pm m)$$

$$\sigma_i(m) = F^{-1}(\frac{1}{2} + m) \qquad \text{for all } i, m$$

First, we can show that this signal structure achieves utility equal to $E(u_S(\theta)\,|\,\theta > \theta_{\frac{1}{2}})$. First remember that, for any random variable x with distribution G, we can define $G(x)$ as another random variable and $G(x) \sim Uniform[0,1]$. Now define $u = u_S(\theta)$ as a random variable with distribution $G(u)$.

Then, we can easily compute the sender's expected utility as follows:

$$E(u_S(a^w)) = \int_0^{0.5} G^{-1}(1-t)dt + \int_{0.5}^{1} G^{-1}(t)dt$$

By a change of variables, we can rewrite the first term to be equal to the second term:

$$\begin{aligned} E(u_S(a^w)) &= 2\int_{0.5}^{1} G^{-1}(t)dt \\ &= 2 \times \int_{u_S(\theta_{\frac{1}{2}})}^{u_S(1)} u dG(u) \\ &= E(u_S(\theta)\,|\,\theta > \theta_{\frac{1}{2}}) \end{aligned}$$

Then we need to show that $E(u_S(\theta)\,|\,\theta > \theta_m)$ is, in fact, an upper bound on the sender's utility. Do this in two steps:

1. Define a subset of F to be a function $G : \Theta \to [0,1]$ s.t. for all $x \geq x'$,

$$G(x) - G(x') \leq F(x) - F(x').$$

Then state that

$$\text{If } G(1) = \frac{1}{2}, \quad \text{then} \quad \frac{\int_0^1 u_S(\theta) dG(\theta)}{\frac{1}{2}} \leq E(u_S(\theta) \,|\, \theta > \theta_m)$$

In other words, take any half of the θ population, and their average is no higher than the average of the top half of the population.

Proof: Define the top half of F to be the subset $F_{\frac{1}{2}}$ such that $F_{\frac{1}{2}}(\theta) \equiv \min\{\, 0 \,,\, F(\theta) - \frac{1}{2}\}$.

($F_{\frac{1}{2}}$ assigns zero probability to all θ below the median of F, but preserves density f for all θ above the median). It is easy to show that, if $G \neq F_{\frac{1}{2}}$ is a subset of F and if $G(1) = \frac{1}{2}$, then $F_{\frac{1}{2}}$ FOSD G. That's because $F_{\frac{1}{2}}$ already assigns the maximum possible probability to high values of θ.

2. Now suppose there exists a signal structure that achieves a utility strictly higher than $E(u_S(\theta) \,|\, \theta > \theta_{\frac{1}{2}})$. Then construct a subset of F which consists of all the realisations of θ above each posterior median. That subset has probability $\frac{1}{2}$ and average higher than u, which contradicts 1.

   Specifically, for each posterior belief F^m corresponding to message (signal realisation) m, define

$$F^m_{\frac{1}{2}}(\theta) \equiv \min\{\, 0 \,,\, F^m(\theta) - \frac{1}{2}\}.$$

   Now the sender's utility is the expectation of the posterior median:

$$u = \int_m u_S(\theta_{\frac{1}{2}}(m)) \Pr(m)$$

Since by definition, $F^m_{\frac{1}{2}}(\theta) = 0$ for all $\theta < \theta_{\frac{1}{2}}(m)$, the above expression is smaller than if we replace $u_S(\theta_{\frac{1}{2}}(m))$ with $\frac{\int_0^1 u_S(\theta)\, dF^m_{\frac{1}{2}}(\theta)}{\frac{1}{2}}$, which is the posterior expectation above $\theta_{\frac{1}{2}}(m)$.

Then we can construct a subset of F by combining all the posterior subsets $F^m_{\frac{1}{2}}$. Define $G(\theta) = \int_m F^m_{\frac{1}{2}}(\theta)\text{Pr }(m)\,dm$. We can show trivially that $G(1) = \frac{1}{2}$. We can also show that G is a subset of F, by using the fact that $F(\theta) = \int_m F^m(\theta)\text{Pr }(m)\,dm$.

$$\frac{\int_0^1 u_S(\theta)\,dG(\theta)}{\frac{1}{2}} = \int\limits_m \frac{\int_0^1 \; u_S(\theta)\; dF^m_{\frac{1}{2}}(\theta)}{\frac{1}{2}} \text{ Pr }(m)\,dm$$

We already established that the RHS was larger than u, which is larger than $E(u_S(\theta) \mid \theta > \theta_m)$. Contradiction.

A.3 Chapter 2: Optimal private signal: Example code in R

```
#Start with number of discrete points
N<-15
#Prior
prior <- rep((1/N),times=N) #Uniform distribution, but we can do
    anything
prior <- p/sum(p) #Just to make sure it adds up to 1, but should not
    matter at all

#The dimensions are s1, s2 and theta
arr <- array(dim=c(N,N,N))

#Winner given s1, s2, theta
w <- arr
for(i in 1:N){
  w[,,i] <- (0.5)*upper.tri(w[,,i], diag = TRUE) +
                                (0.5)*upper.tri(w[,,i], diag = FALSE)
}
for(i in 1:N){
  for(j in 1:N){
    for(k in 0:min(N,max(2*i-j,0))){
      w[j,k,i] <- 1-w[j,k,i]
    }}
}

#Tie when equidistant from theta
for(i in 1:N){
  for(j in max(2*i-N,0):i){
    w[j,(2*i-j),i] <- 0.5
    w[(2*i-j),j,i] <- 0.5
  }
}

#Expected winning policy
#First state the policies of each candidate
a1 <- arr
for(i in 1:N){
  a1[i,,] <- (i-1)/(N-1)
}
a2 <- arr
for(i in 1:N){
```

```
  a2[,i,] <- (i-1)/(N-1)
}

#Expected winning policy
aw <- w*a1 +(1-w)*a2

#Sender's utility is sum(arr*aw), where arr is joint dist.
#If u_S is not linear, we need to apply it to all elements of a1, a2
#So our coefficients in objective function are (aw) in vector form
aw <- as.vector(aw)
#Then need constraint matrix, including the fact that each marginal
       #of theta has to add up to f(theta)

#Then we can use lpSolve to solve it

#Constraint matrix
#First, for each theta, the arr elements for that theta add up to
      #f(theta), which could be 1/N if uniform.
#Then, we need incentive compatibility for each action for j

#First, I need a vector that's 1 only for one theta, 0 for others
C1 <- matrix(0,nrow=N,ncol=(N^3))
for(n in 1:N){
  C1[n,((n-1)*(N^2)+1):(n*(N^2))] <- 1
}

#Now rbind with obedience constraints
#Obedience constraint: Given each action recommended, it's not
         #better to do another action
#An action is a row for player 1, a column for player 2

C2 <- matrix(nrow=(N^2),ncol=(N^3))
for(n in 1:N){ #recommended actions
  for(i in 1:N){ #deviations
    d <- array(0,dim=c(N,N,N))
    d[,n,] <- 1-w[,n,] - (1-w[,i,])
    C2[((n-1)*N+i),] <- as.vector(d)
  }}
C3 <- matrix(nrow=(N^2),ncol=(N^3))
for(n in 1:N){ #recommended actions
  for(i in 1:N){ #deviations
    d <- array(0,dim=c(N,N,N))
```

```
    d[n,,] <- w[n,,] - w[i,,]
    C3[((n-1)*N+i),] <- as.vector(d)
  }}

C <- rbind(C1,C2,C3)
C <- t(C)
#Apparently transposing the constraints matrix first makes it quicker
#Now what are the RHS values for the constraints
#For C1, it's just f(theta)
#For C2 and C3, it's 0

cons <- rep(0, times=(N+2*N^2))
cons[1:N] <- prior

# Load lpSolve
require(lpSolve)

#Coefficients of obj fn: aw
#Constraints: C
#Constraints RHS: cons

# Direction of the constraints
cons_direction  <- vector(length=(N+2*N^2))
cons_direction[1:N] <- "="
cons_direction[(N+1):(N+2*N^2)] <- ">="

# Find the optimal solution
optimum <-  lp(direction="max",
               objective.in = aw,
               const.mat = C,
               const.dir = cons_direction,
               const.rhs = cons,
               transpose.constraints=FALSE)

# Print status: 0 = success, 2 = no feasible solution
print(optimum$status)

# Display arg max
best_sol <- optimum$solution
print(best_sol)

# Check the value of objective function at optimal point
```

```
print(paste("Sender's utility ", optimum$objval, sep=""))

#Check that all non-diagonal elements (private messgaes)
#of the arg max are zero
offdiag <- array(1,dim=c(N,N,N))
for(i in 1:N){
  offdiag[i,i,] <- 0
}
anyprivatem <- max(best_sol*as.vector(offdiag))
print(anyprivatem)
```

A.4 Chapter 3: Additional Figures and Tables

A.4.1 Figures

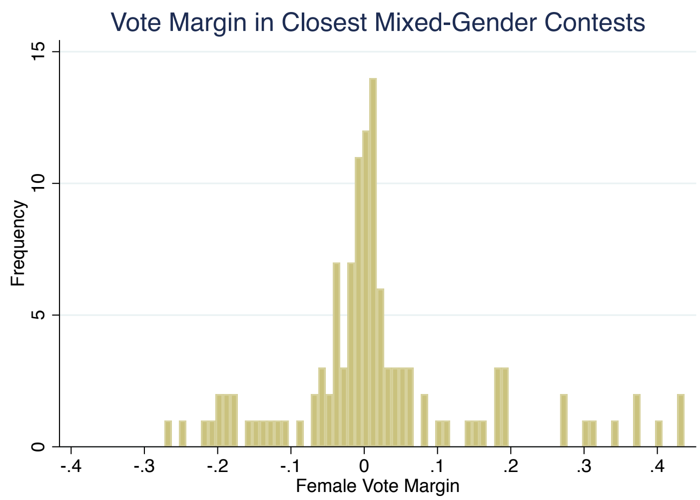


Figure A·1: Histogram of Vote Margins in Closest Race in Each City

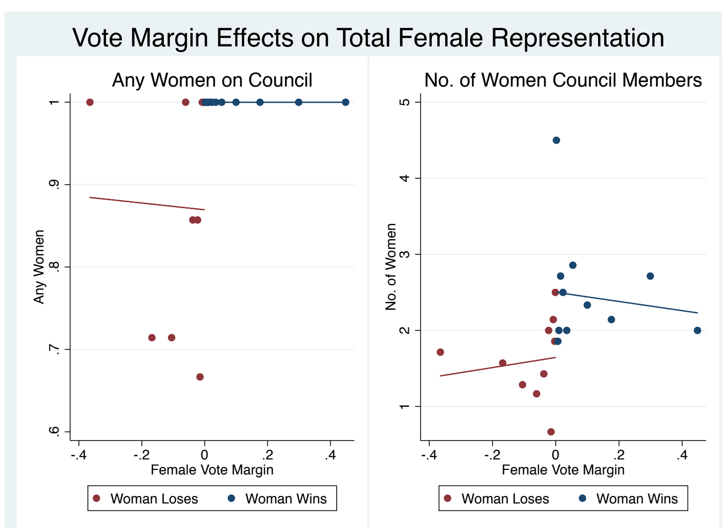


Figure A·2: Female Representation is Discontinuous in the Running Variable

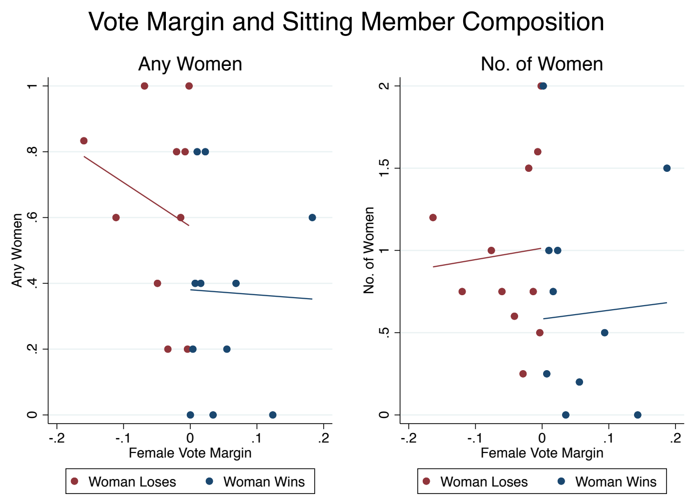


Figure A·3: Continuity Check: Gender Composition of Sitting Members of the Council

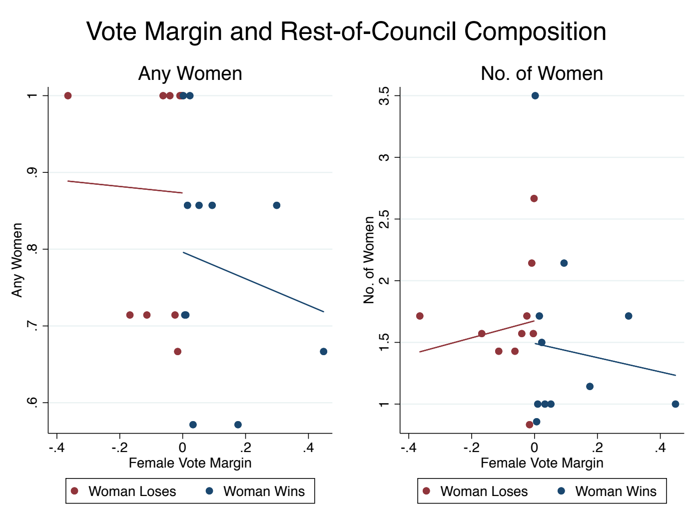


Figure A·4: Continuity Check: Gender Composition of the Rest of the Council

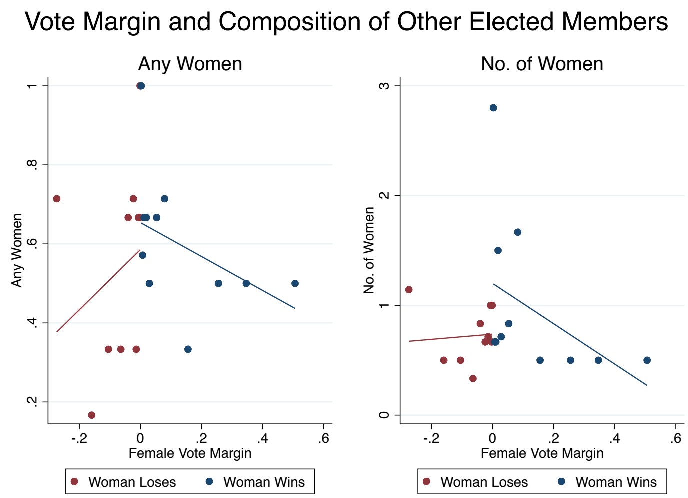


Figure A·5: Continuity Check: Gender Composition of Other Elected Members

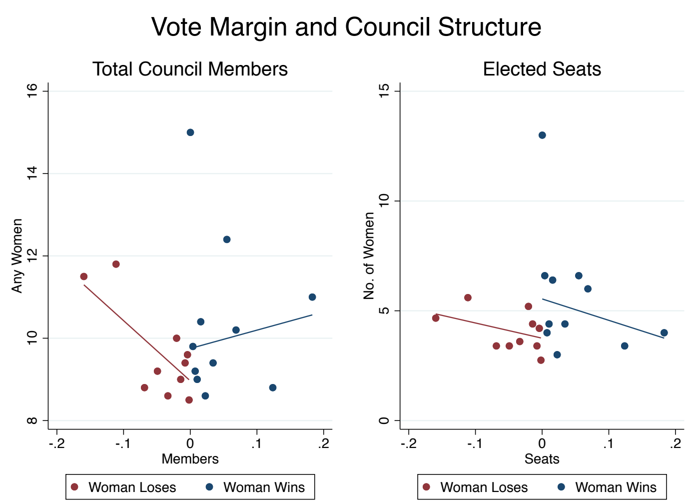


Figure A·6: Continuity Check: Council Size and Number of Elected Seats

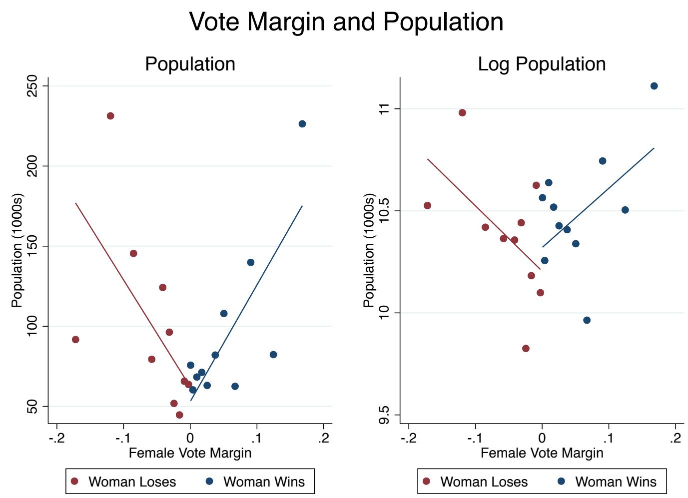


Figure A·7: Continuity Check: Population

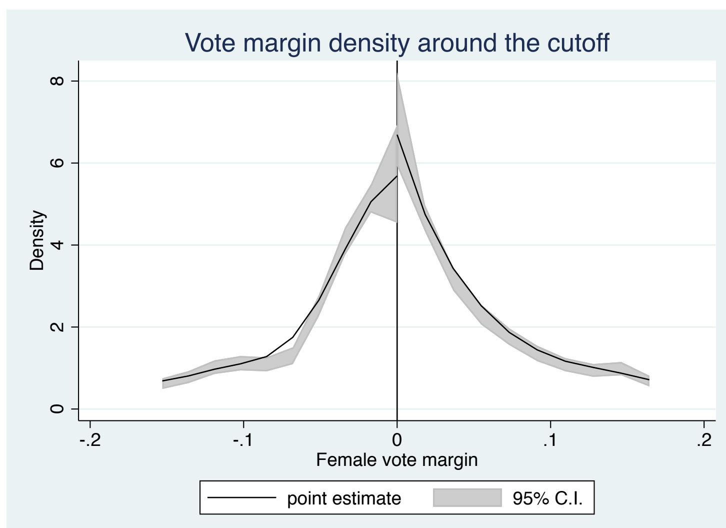


Figure A·8: McCrary Test: Density of Vote Margin Around Cutoff

A.4.2 Tables

Table A.1: Panel Specification: Real and Log Total Expenditure Correlations with Female Representation

	Real Total Expenditure ($)					Log Total Expenditure				
Any Woman on Council	-7956847.0		-8860863.3		-5574080.2	-0.0177		-0.0262		-0.0235
	(9616089.5)		(11565532.0)		(9742713.1)	(0.0190)		(0.0229)		(0.0193)
Percent of Council Women		-7484174.1	3698156.1				0.00152	0.0346		
		(21850868.6)	(26278358.4)				(0.0432)	(0.0520)		
No. of Women on Council				-5856336.3					0.00646	
				(3986492.7)					(0.00791)	
2+ Women on Council					2897045.4					0.0205
					(7997305.2)					(0.0159)
3+ Women on Council					-11182897.8					0.00993
					(10027189.4)					(0.0199)
Observations	3953	3953	3953	3953	3953	3953	3953	3953	3953	3953

Standard errors are clustered at the city level and are in parentheses. All regressions include year and city fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.2: Panel Specification: Real and Log Per Capita Expenditure Correlations with Female Representation

	Real Per Capita Expenditure ($)					Log Per Capita Expenditure				
Any Woman on Council	-1007.5		1115.3		-811.8	-0.0189		-0.0304		-0.0261
	(1236.0)		(1485.0)		(1256.3)	(0.0187)		(0.0225)		(0.0190)
Percent of Council Women		-7276.3***	-8683.8**				0.00861	0.0470		
		(2805.7)	(3374.2)				(0.0426)	(0.0512)		
No. of Women on Council				-433.4					0.00751	
				(514.2)					(0.00780)	
2+ Women on Council					-1085.1					0.0258*
					(1031.2)					(0.0156)
3+ Women on Council					496.7					0.0111
					(1292.9)					(0.0196)
Observations	3953	3953	3953	3953	3953	3953	3953	3953	3953	3953

Standard errors are clustered at the city level and are in parentheses. All regressions include year and city fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Panel Specification: Shares of Expenditure Categories Correlations with Female Representation

	Share of Expenditure on Public Utility					Share of Expenditure on Health and Hospitals				
Any Woman on Council	-0.00202 (0.00443)		0.000214 (0.00533)		-0.00112 (0.00450)	-0.00654*** (0.00149)		-0.00638*** (0.00178)		-0.00650*** (0.00151)
Percent of Council Women		-0.00889 (0.0101)	-0.00916 (0.0121)				-0.00888** (0.00347)	-0.000696 (0.00415)		
No. of Women on Council				-0.000735 (0.00184)					-0.00147** (0.000619)	
2+ Women on Council					-0.00143 (0.00370)					-0.000675 (0.00125)
3+ Women on Council					-0.000758 (0.00463)					0.000474 (0.00155)
Observations	3952	3952	3952	3952	3952	3487	3487	3487	3487	3487
	Share of Expenditure on Parks and Recreation					Share of Expenditure on Libraries				
Any Woman on Council	0.000353 (0.00193)		-0.000265 (0.00232)		0.000475 (0.00196)	-0.000526 (0.000832)		-0.00137 (0.000997)		-0.000491 (0.000844)
Percent of Council Women		0.00220 (0.00439)	0.00253 (0.00527)				0.00178 (0.00191)	0.00352 (0.00229)		
No. of Women on Council				0.000486 (0.000801)					0.000237 (0.000344)	
2+ Women on Council					-0.00126 (0.00161)					0.000509 (0.000691)
3+ Women on Council					0.00337* (0.00201)					-0.000252 (0.000863)
Observations	3923	3923	3923	3923	3923	3658	3658	3658	3658	3658
	Share of Expenditure on Housing and Community Development					Share of Expenditure on Airports and Water Ports				
Any Woman on Council	-0.00344 (0.00290)		-0.00140 (0.00349)		-0.00263 (0.00295)	-0.000451 (0.000723)		-0.00110 (0.000865)		-0.000603 (0.000733)
Percent of Council Women		-0.0101 (0.00659)	-0.00831 (0.00792)				0.00134 (0.00168)	0.00275 (0.00201)		
No. of Women on Council				-0.00172 (0.00121)					0.000219 (0.000299)	
2+ Women on Council					-0.00353 (0.00242)					0.000735 (0.000604)
3+ Women on Council					-0.00119 (0.00303)					0.000332 (0.000747)
Observations	3953	3953	3953	3953	3953	3475	3475	3475	3475	3475
	Share of Expenditure on Police and Fire					Share of Expenditure on Sewerage and Waste				
Any Woman on Council	0.000983 (0.00384)		-0.00473 (0.00461)		-0.000472 (0.00390)	-0.00423 (0.00336)		-0.00278 (0.00402)		-0.00381 (0.00340)
Percent of Council Women		0.0175** (0.00872)	0.0235** (0.0105)				-0.00968 (0.00782)	-0.00611 (0.00937)		
No. of Women on Council				0.00239 (0.00159)					-0.00162 (0.00139)	
2+ Women on Council					0.00599* (0.00319)					-0.00216 (0.00282)
3+ Women on Council					0.00237 (0.00399)					-0.000383 (0.00349)
Observations	3837	3837	3837	3837	3837	3487	3487	3487	3487	3487
	Share of Expenditure on Corrections					Share of Expenditure on Roads and Parking				
Any Woman on Council	0.0000475 (0.000752)		-0.000108 (0.000870)		0.0000655 (0.000770)	0.00862** (0.00415)		0.0156*** (0.00497)		0.00936** (0.00421)
Percent of Council Women		0.000647 (0.00191)	0.000784 (0.00221)				-0.00956 (0.00961)	-0.0294** (0.0115)		
No. of Women on Council				0.0000987 (0.000253)					-0.00156 (0.00172)	
2+ Women on Council					0.0000500 (0.000600)					-0.00340 (0.00345)
3+ Women on Council					0.0000851 (0.000631)					-0.00839* (0.00429)
Observations	1161	1161	1161	1161	1161	3535	3535	3535	3535	3535

Standard errors are clustered at the city level and are in parentheses. All regressions include year and city fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: RD Estimates: Effect of Close Female Win on Expenditure, Without Controls

	Total Real Expenditure			Total Log Expenditure		
	Optimal Bandwidth	Half Optimal Bandwidth	Twice Optimal Bandwidth	Optimal Bandwidth	Half Optimal Bandwidth	Twice Optimal Bandwidth
Female Winner	6403567.3	6699025.7	116626475.5	0.0539	0.0378	-0.0716
	(28511351.1)	(39251614.8)	(104797748.5)	(0.159)	(0.186)	(0.140)
Bandwidth	0.0670	0.0335	0.134	0.160	0.0800	0.320

	Total Expenditure Per Capita			Log Total Expenditure Per Capita		
	Optimal Bandwidth	Half Optimal Bandwidth	Twice Optimal Bandwidth	Optimal Bandwidth	Half Optimal Bandwidth	Twice Optimal Bandwidth
Female Winner	-443.6	-644.3*	-670.6*	-0.0926	-0.204**	-0.124**
	(302.5)	(362.8)	(354.4)	(0.0687)	(0.0844)	(0.0558)
Bandwidth	0.123	0.0615	0.246	0.148	0.0740	0.296

Standard errors are clustered at the city level and are in parentheses. All regressions include year and state fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: RD Polynomial Estimates: Effect of Close Female Win on Expenditure, Controlling for Population and Winner's Experience

	Total Real Expenditure			Total Log Expenditure		
Female Winner	-6814303.8	-2438897.7	-4127122.1	-0.0119	-0.272	-0.609**
	(20450422.4)	(23923400.0)	(27330327.9)	(0.178)	(0.239)	(0.289)
Polynomial	2	3	4	2	3	4

	Total Expenditure Per Capita			Log Total Expenditure Per Capita		
Female Winner	-555.6*	-620.7*	-498.1	-0.166*	-0.258**	-0.307**
	(323.7)	(339.4)	(310.7)	(0.0937)	(0.110)	(0.125)
Polynomial	2	3	4	2	3	4

Standard errors are clustered at the city level and are in parentheses.
All regressions include year and state fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.6: RD Estimates: Effect of Close Female Win on Expenditure Composition

Bandwidth	0.100	0.150	0.200	0.250
Dependent Variable:				
% Public Utility Expenditure	-0.0162	-0.0154	-0.0138	-0.0136
	(0.0127)	(0.0116)	(0.0102)	(0.00965)
% Health and Hospital Expenditure	0.00415	0.00709	0.00524	0.00524
	(0.0139)	(0.0120)	(0.0112)	(0.00985)
% Parks and Recreation Expenditure	0.000136	0.00327	0.00377	0.00203
	(0.00631)	(0.00545)	(0.00489)	(0.00452)
% Library Expenditure	-0.00104	-0.000806	-0.00246	-0.00199
	(0.00230)	(0.00193)	(0.00168)	(0.00151)
% Housing and Community Dev Expenditure	-0.00202	-0.00503	-0.00492	-0.00236
	(0.00779)	(0.00676)	(0.00613)	(0.00593)
% Airports and Water Ports Expenditure	-0.000388	-0.000562	-0.00697*	-0.00517
	(0.00282)	(0.00324)	(0.00372)	(0.00411)
% Police and Fire Expenditure	0.0109	0.00657	0.00680	0.00981
	(0.0127)	(0.0112)	(0.0100)	(0.00910)
% Sewerage and Waste Expenditure	0.00750	0.00362	0.00406	0.00461
	(0.0127)	(0.0116)	(0.0104)	(0.00927)
% Correctional Expenditure	0.000175	-0.000935	-0.000680	-0.000283
	(0.00139)	(0.000682)	(0.000658)	(0.000679)
% Roads and Parking Expenditure	0.0122	0.0110	0.0114	0.00757
	(0.00885)	(0.00838)	(0.00767)	(0.00728)
Joint test	8.155	8.479	11.63	8.180
p-value	0.614	0.582	0.311	0.611

Standard errors are clustered at the city level and are in parentheses.
All regressions include year and state fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.7: RD Estimates: Effect of Close Female Win on Log per Capita Expenditure Categories, Controlling for Population and Winner's Experience

Bandwidth	0.100	0.150	0.200	0.250
Dependent Variable:				
Log Per Capita Public Utility Expenditure	-0.272	-0.232	-0.267	-0.316
	(0.306)	(0.271)	(0.243)	(0.221)
Log Per Capita Health and Hospital Expenditure	-0.159	-0.0525	-0.00181	-0.0370
	(0.216)	(0.184)	(0.165)	(0.152)
Log Per Capita Parks and Recreation Expenditure	-0.155	-0.0728	-0.0484	-0.0836
	(0.131)	(0.124)	(0.115)	(0.107)
Log Per Capita Library Expenditure	-0.240	-0.162	-0.282*	-0.321**
	(0.216)	(0.189)	(0.168)	(0.157)
Log Per Capita Housing and Community Dev Expenditure	-0.0964	-0.0275	-0.0364	-0.0364
	(0.124)	(0.111)	(0.0994)	(0.0966)
Log Per Capita Airports and Water Ports Expenditure	0.0585	-0.0523	-0.144	-0.164
	(0.170)	(0.149)	(0.144)	(0.146)
Log Per Capita Police and Fire Expenditure	-0.0372	-0.0206	-0.0659	-0.0591
	(0.0890)	(0.0759)	(0.0629)	(0.0592)
Log Per Capita Sewerage and Waste Expenditure	-0.0747	-0.0814	-0.164	-0.167
	(0.294)	(0.251)	(0.225)	(0.205)
Log Per Capita Correctional Expenditure	-0.0437	-0.121	-0.0872	-0.0361
	(0.159)	(0.1000)	(0.0908)	(0.0790)
Log Per Capita Roads and Parking Expenditure	-0.0561	-0.00188	-0.0225	-0.0456
	(0.0895)	(0.0810)	(0.0762)	(0.0726)
Joint test	3.700	2.814	5.081	6.199
p-value	0.960	0.985	0.886	0.798

Standard errors are clustered at the city level and are in parentheses.
All regressions include year and state fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.8: Continuity Checks: Council and Election Variables

	Council Size	Elected Seats	Rest of Council		Sitting Members		Other Elected Members	
			Any Woman	% Women	Any Woman	% Women	Any Woman	% Women
Female Winner	0.436	0.310	0.0164	-0.0118	0.0356	-0.0605	-0.0588	-0.0402
	(0.720)	(0.744)	(0.119)	(0.0459)	(0.161)	(0.0806)	(0.173)	(0.106)
Bandwidth	0.200	0.200	0.200	0.200	0.200	0.200	0.200	0.200

Standard errors are clustered at the city level and are in parentheses. All regressions include year and state fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A.9: Continuity Check: Population

	Population
Female Winner	-15974.1
	(18388.0)
Bandwidth	0.200

Standard errors are clustered at the city level and are in parentheses.
Regression includes year and state fixed effects.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

References

Alesina, A., Baqir, R., and Easterly, W. (1999). Public Goods and Ethnic Divisions. *The Quarterly Journal of Economics*, 114(4):1243–1284.

Alesina, A. and La Ferrara, E. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, 43(3):762–800.

Austen-Smith, D. (1993). Information and Influence: Lobbying for Agendas and Votes. *American Journal of Political Science*, 37(3):799.

Bagues, M. F. and Esteve-Volart, B. (2010). Can gender parity break the glass ceiling? evidence from a repeated randomized experiment. *The Review of Economic Studies*, 77(4):1301–1328.

Beach, B. and Jones, D. B. (2017). Gridlock: Ethnic diversity in government and the provision of public goods. *American Economic Journal: Economic Policy*, 9(1):112–36.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., and Topalova, P. (2009). Powerful women: Does exposure reduce bias? *The Quarterly Journal of Economics*, 124(4):1497–1540.

Beaman, L., Duflo, E., Pande, R., and Topalova, P. (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in india. *Science*, 335(6068):582–586.

Bergemann, D. and Morris, S. (2019). Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95.

Besley, T. and Coate, S. (1997). An economic model of representative democracy. *Quarterly Journal of Economics*, 112(1):85–114.

Callander, S. (2005). Electoral competition in heterogeneous districts. *Journal of Political Economy*, 113(5):1116–1145.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Chakraborty, A. and Harbaugh, R. (2007). Comparative cheap talk. *Journal of Economic Theory*, 132(1):70–94.

Chakraborty, A. and Harbaugh, R. (2010). Persuasion by Cheap Talk. *The American Economic Review*, 100(5):2361–2382.

Chattopadhyay, R. and Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica*, 72(5):1409–1443.

Coffman, K. B. (2014). Evidence on self-stereotyping and the contribution of ideas. *The Quarterly Journal of Economics*, 129(4):1625–1660.

Crawford, V. P. and Sobel, J. (1982). Strategic Information Transmission. *Econometrica*, 50(6):1431.

Downs, A. (1957). An Economic Theory of Political Action in a Democracy. *Journal of Political Economy*, 65(2):135–150.

Farrell, J. and Gibbons, R. (1989). Cheap Talk with Two Audiences. *The American Economic Review*, 79:1214–1223.

Ferreira, F. and Gyourko, J. (2014). Does gender matter for political leadership? the case of u.s. mayors. *Journal of Public Economics*, 112:24–39.

Friedman, J. N. and Holden, R. T. (2008). Optimal gerrymandering: Sometimes pack, but never crack. *American Economic Review*, 98(1):113–144.

Gentzkow, M. and Kamenica, E. (2016). A Rothschild-Stiglitz Approach to Bayesian Persuasion. *American Economic Review*, 106:597–601.

Gentzkow, M. and Kamenica, E. (2017). Disclosure of endogenous information. *Economic Theory Bulletin*, 5(1):47–56.

Gitmez, A. and Molavi, P. (2018). Media Capture: A Bayesian Persuasion Approach. Available from http://economics.mit.edu/files/17030

Grossman, G., Helpman, E., Grossman, G., and Helpman, E. (1994). Protection for Sale. *American Economic Review*, 84(4):833–50.

Grossman, G. M. and Helpman, E. (1996). Electoral Competition and Special Interest Politics. *The Review of Economic Studies*, 63(2):265.

Grossman, G. M. and Helpman, E. (2001). *Special interest politics*. MIT.

Heese, C. and Lauermann, S. (2019). The Bayes Correlated Equilibria of Large Elections. *Working Paper*, (April). https://www.aeaweb.org/conference/2019/preliminary/paper/Gz4nsfEQ

Herbst, S. (1998). *Reading Public Opinion: How Political Actors View the Democratic Process*. University of Chicago Press.

Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, 39(153):41.

Kamada, Y. and Kojima, F. (2014). Voter preferences, polarization, and electoral policies. *American Economic Journal: Microeconomics*, 6(4):203–236.

Kamenica, E. (2019). Bayesian Persuasion and Information Design. *Annual Review of Economics*, 11(1).

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

Kartik, N. and Mcafee, P. R. (2007). Signaling character in electoral competition. *American Economic Review*, 97(3):852–870.

Kartik, N. and Van Weelden, R. (2019). Informative cheap talk in elections. *Review of Economic Studies*, 86(2):755–784.

Kolotilin, A. and Zapechelnyuk, A. (2019). Persuasion Meets Delegation. *Discussion Paper Series, School of Economics and Finance.* https://arxiv.org/abs/1902.02628

Lipnowski, E. and Ravid, D. (2017). Cheap Talk with Transparent Motives. *Ssrn*, (May):1–56.

Potters, J. and van Winden, F. (1992). Lobbying and asymmetric information. *Public Choice*, 74(3):269–292.

Ray, D. and Esteban, J. (2017). Conflict and Development. *Annual Review of Economics*, 9(1):263–293.

Roemer, J. E. (1994). A theory of policy differentiation in single issue electoral politics. *Social Choice and Welfare*, 11:355–380.

Curriculum Vitae

[redacted]